

# **Interaktionsverhalten von frühpädagogischen Fachkräften**

**Qualitätsentwicklung durch  
online-basiertes Videofeedback**

Inauguraldissertation  
zur Erlangung des Doktorgrades  
der Philosophie an der Ludwig-Maximilians-Universität  
München

vorgelegt von

**Nesiré Kappauf**

(geb. Schauland)

München, 2018

Erstgutachter: Prof. Dr. Rudolf Tippelt

Zweitgutachterin: Prof. Dr. Fabienne Becker-Stoll

Datum der mündlichen Prüfung: 15.05.2018



# Danksagung

Ich möchte an dieser Stelle meinen Dank an alle aussprechen, die mich während meiner Promotionszeit unterstützt haben.

Bei Prof. Dr. Rudolf Tippelt und Prof. Dr. Fabienne Becker-Stoll möchte ich mich für die fachliche Betreuung und Begleitung der Arbeit und das Vertrauen, dass Sie mir geschenkt haben, bedanken.

Ein besonderer Dank gilt den pädagogischen Fachkräften und Leitungen sowie den Eltern und Kindern, ohne deren Teilnahme und Engagement dieses Projekt nicht möglich gewesen wäre.

Desweiteren möchte ich herzlich meinen Kolleginnen und Kollegen am Staatsinstitut für Frühpädagogik, allen voran den MitarbeiterInnen aus den „BiSS“-Evaluationsprojekten, für das offene Ohr und die motivierenden Worte danken. Insbesondere die Projektleitung Frau Dr. Claudia Wirts, meine ProjektkollegInnen Sina Fischer und Dr. Erik Danay sowie die wissenschaftlichen Hilfskräfte Nicole Baginski und Myriam Schnider haben durch ihre Unterstützung zum Gelingen dieses Projekts beigetragen.

Schließlich gilt ein großer Dank meiner Familie und meinen Freunden, die immer an mich geglaubt haben und unterstützend an meiner Seite standen. Ganz besonders möchte ich meinem Freund Hannes danken, der mich durch alle Höhen und Tiefen meiner Promotionszeit begleitet und mir die nötige Kraft gegeben hat, diese Herausforderung zu meistern.

# Zusammenfassung

Vor dem Hintergrund von empirischen Belegen für eine niedrige bis mittlere Qualität sprachlicher Anregung in Kindertageseinrichtungen wurde von der Autorin im Rahmen des Projekts „BiSS-Evaluation“ ein Konzept entwickelt, mit dem frühpädagogischen Fachkräften Feedback zu ihrem sprachlichen Interaktionsverhalten gegeben werden kann. Dabei werden gemeinsam mit der Fachkraft Videosequenzen aus ihrem pädagogischen Alltag hinsichtlich ihres sprachförderlichen Verhaltens online mithilfe einer Videokonferenzsoftware reflektiert. Ziel der vorliegenden Arbeit war es, die Effekte eines auf dem Konzept basierenden Feedbackgesprächs auf das sprachliche Interaktionsverhalten von Fachkräften während einer Bilderbuchbetrachtung zu überprüfen. An einer zufällig ausgewählten Stichprobe von 96 Fachkräften, die mit Kindern zwischen drei und sechs Jahren arbeiten, wurde die Evaluation des Feedbacks in einem Prä-Post-Kontrollgruppendesign durchgeführt. Bei allen teilnehmenden Fachkräften wurde im Abstand von ca. vier Wochen jeweils eine Videoaufnahme während einer Bilderbuchbetrachtung angefertigt, welche hinsichtlich der Qualität der sprachlichen Anregung sowie in Bezug auf die Häufigkeit verschiedener sprachlicher Verhaltensweisen ausgewertet wurde. Die Fachkräfte in der Feedbackbedingung nahmen zwischen den zwei Messzeitpunkten an einem ca. einstündigen online-basierten Feedbackgespräch teil, bei dem einzelne Sequenzen aus der ersten Videoaufnahme reflektiert wurden. Die Durchführung der Feedbackgespräche hat insgesamt gut funktioniert, wobei es aufgrund einer mangelhaften Ausstattung in den Kitas zum Teil noch zu technischen Problemen gekommen ist. Die teilnehmenden Fachkräfte bewerteten das Feedback sehr positiv. Die Ergebnisse auf Performanz-Ebene deuten darauf hin, dass sich die Fachkräfte aus der Feedbackgruppe ( $N = 48$ ) signifikant in ihrem sprachlichen Interaktionsverhalten verbessert haben, während in der Kontrollgruppe ( $N = 48$ ) keine bedeutenden Veränderungen festgestellt wurden. Bei den positiven Veränderungen in der Feedbackgruppe handelt es sich insgesamt um mittlere bis große Effekte. Auf individueller Ebene fanden die meisten positiven Veränderungen in der globalen Einschätzung der sprachlichen Anregung sowie bei den Fragen mit Antwortmöglichkeit für die Kinder und den nicht korrektiven Wiederholungen kindlicher Äußerungen statt. Die Veränderungen innerhalb der Feedbackgruppe konnten nicht durch Charakteristika der Person, wie die Motivation, die Einstellung gegenüber Feedback oder die Zielorientierung der Fachkräfte vorhergesagt werden. Es kann festgehalten werden, dass online-basiertes Videofeedback als ressourcenschonendes und effektives Format zur Verbesserung der sprachlichen Interaktionsqualität von frühpädagogischen Fachkräften dienen kann.

# Inhaltsverzeichnis

1	Problemstellung, Ziele und Aufbau dieser Arbeit.....	8
2	Pädagogisches Handeln von frühpädagogischen Fachkräften.....	11
2.1	Interaktionsqualität als Kernelement von Prozessqualität.....	11
2.1.1	Bedeutung von Interaktionen für die kindliche Entwicklung .....	11
2.1.2	Operationalisierung von Fachkraft-Kind-Interaktionen .....	17
2.1.3	Empirische Befunde zur Interaktionsqualität und ihrem Zusammenhang zur kindlichen Entwicklung .....	22
2.2	Möglichkeiten der Qualitätsentwicklung.....	27
2.2.1	Ansätze zur Verbesserung der Prozessqualität.....	28
2.2.2	Fort-und Weiterbildung als Säule der Qualitätsentwicklung .....	29
3	Feedback.....	38
3.1	Begriffsbestimmung.....	38
3.1.1	Definition von Feedback.....	38
3.1.2	Inhalte und Formen von Feedback .....	39
3.2	Wirkungen und Bedingungen von Feedback.....	41
3.2.1	Funktionsweisen von Feedback.....	41
3.2.2	Empirische Befunde zu den Funktionsweisen von Feedback .....	46
3.2.3	Zusammenfassende Ergebnisse von Meta-Analysen .....	49
3.2.4	Bedingungsfaktoren für die Effekte von Feedback.....	50
3.2.5	Theoretische Erklärungsansätze für die Wirkung von Feedback.....	58
3.2.6	Feedback im (früh-) pädagogischen Kontext.....	64
3.3	Implikationen für die Gestaltung von Feedback .....	66
4	Zusammenfassung, Forschungslücke und Ziele.....	70
5	Konzept für ein online-basiertes Videofeedback .....	72
5.1	Eckpfeiler des Feedback-Konzepts.....	72
5.2	Vorgehen beim Feedbackgespräch .....	75

5.2.1	Vorbereitung des Feedbackgesprächs.....	75
5.2.2	Durchführung des Feedbackgesprächs .....	78
5.2.3	Nachbereitung des Feedbackgesprächs.....	83
5.3	Das Feedback-Konzept auf einen Blick .....	84
5.4	Erprobung des Feedback-Konzepts .....	84
6	Umsetzung und Überprüfung des Feedback-Konzepts: Fragestellung und Methode.....	87
6.1	Fragestellungen und Hypothesen.....	87
6.2	Design.....	89
6.3	Variablen und Instrumente .....	90
6.3.1	Abhängige Variablen .....	90
6.3.2	Intervention .....	101
6.3.3	Einfluss- und Kontrollvariablen .....	105
6.4	Stichprobe.....	113
6.5	Ablauf der Studie.....	119
6.6	Statistische Auswertung.....	120
6.6.1	Beurteilerübereinstimmung und -reliabilität.....	121
6.6.2	Variablenbildung.....	123
6.6.3	Überprüfung der Hypothesen und Fragestellungen .....	124
7	Umsetzung und Überprüfung des Feedback-Konzepts: Ergebnisse .....	132
7.1	Umsetzung der Feedbackgespräche.....	132
7.2	Bewertung des Feedbackgesprächs durch die Fachkräfte.....	135
7.3	Effekte des Feedbackgesprächs auf das sprachliche Interaktionsverhalten.....	137
7.3.1	Effekte auf das „Language Modeling“ der Fachkräfte .....	137
7.3.2	Effekte auf die Anzahl sprachlicher Verhaltensweisen.....	141
7.3.3	Individuelle Veränderungen durch das Feedback.....	162
7.4	Einfluss von anderen Variablen auf die Effekte des Feedbacks .....	164
7.4.1	Kontrolle von Störvariablen.....	165
7.4.2	Vorhersage der Feedbackwirkung durch verschiedene Einflussfaktoren.....	167

8	Diskussion der Ergebnisse und Ableitung von Implikationen.....	171
8.1	Erprobung des Feedback-Konzepts .....	171
8.2	Effekte des Feedbackgesprächs.....	177
8.3	Grenzen der vorliegenden Studie.....	189
9	Fazit und Ausblick .....	191
10	Literaturverzeichnis .....	195
11	Tabellenverzeichnis.....	215
12	Abbildungsverzeichnis .....	218
13	Abkürzungsverzeichnis.....	220
14	Anhang .....	221

# 1 Problemstellung, Ziele und Aufbau dieser Arbeit

Die Berufstätigkeit von Müttern hat in den letzten Jahrzehnten stark zugenommen (Hofer, Wild & Noack, 2002). Immer mehr Familien sind daher auf eine Betreuung außerhalb der Familie angewiesen. In den letzten Jahren erfolgte ein enormer quantitativer Ausbau der Betreuungsplätze. Mittlerweile werden deutschlandweit bereits über 90 % der Kinder ab drei Jahren in Kindertageseinrichtungen betreut (Statistisches Bundesamt, 2016). Aus diesem Grund hat die außerfamiliäre Kindertagesbetreuung als Kontext für die kindliche Entwicklung enorm an Bedeutung gewonnen. In zahlreichen Stellungnahmen wird deshalb betont, dass die Qualitätssicherung innerhalb der Kindertagesbetreuung eine bedeutende Rolle spielt (z.B. Deutsche Liga für das Kind, 2008; Deutsches Jugendinstitut, 2008).

Trotz der zunehmenden Bedeutung, die dem Begriff „Qualität“ im Bildungswesen beigegeben wird, gibt es eine wirklich präzise Definition bislang nicht (Helmke, Hornstein & Terhart, 2000). Was unter einer guten Qualität verstanden wird, hängt von der jeweiligen Interessengruppe und ihren Zielen ab. So können Eltern unter einer hohen Qualität in der Kita etwas ganz anderes verstehen als Verantwortliche in der Politik. Zusätzlich zu den unterschiedlichen Perspektiven auf die Bestimmung von Qualität, werden verschiedene Dimensionen von Qualität unterschieden. Jede Dimension von Qualität lässt sich wiederum von mehreren Indikatoren genauer beschreiben. Nach Tietze et al. (1998) kann im vorschulischen Bereich zwischen Struktur-, Orientierungs- und Prozessqualität unterschieden werden. Strukturqualität meint die *„situationsunabhängigen, zeitlich stabilen Rahmenbedingungen in den Betreuungsangeboten für Kinder“* (Wissenschaftlicher Beirat für Familienfragen beim BMFSFJ, 2008, S. 13). Zur strukturellen Dimension gehören zum Beispiel die Gruppengröße, der Betreuungsschlüssel, die Qualifikation der Fachkräfte oder die räumliche Ausstattung. Die Strukturqualität stellt die Dimension dar, welche direkt von der politischen Steuerung abhängig ist, vor allem in Form von finanziellen Investitionen. Unter Orientierungsqualität werden die Werte, Einstellungen, Ziele und das Wissen der pädagogischen Mitarbeiter, welche Grundlage ihrer pädagogischen Arbeit sind, subsumiert. Die Prozessqualität bezeichnet die Gesamtheit der täglichen Erfahrungen, die das Kind mit der räumlich-materiellen und sozialen Umwelt in der Kita macht. Wesentliches Merkmal der Prozessqualität sind daher die Interaktionen zwischen pädagogischen Fachkräften<sup>1</sup> und Kindern. Eine Vielzahl von Theorien und wissenschaftlichen Untersuchungen hebt die Bedeutung

---

<sup>1</sup> In Kindertageseinrichtungen sind ErzieherInnen, aber auch andere Berufsgruppen (z.B. KinderpflegerInnen) als Fach- oder Ergänzungskräfte tätig. In dieser Arbeit sind mit dem Begriff der frühpädagogischen Fachkraft alle pädagogisch tätigen Personen gemeint, da sie zur Qualität der Betreuung beitragen. Da die Mehrheit der Fachkräfte weiblich ist, wird bei entsprechenden Bezeichnungen die weibliche Berufsform verwendet.

von Interaktionen für kindliche Entwicklungs- und Bildungsprozesse hervor. In der Interaktion mit Erwachsenen bauen Kinder Beziehungen oder Bindungen auf, erwerben wichtige kognitive und sozial-emotionale Fähigkeiten und erlernen die Sprache (Siegler, DeLoache & Eisenberg, 2008). Vor allem hinsichtlich lernunterstützender Interaktionen deuten die Ergebnisse empirischer Studien jedoch auf ein eher niedriges bis mittleres Qualitätsniveau hin (z.B. Hamre et al., 2013; Wirts, Wertfein & Wildgruber, 2017). Wenn Kinder im Kontext von außerfamiliärer Betreuung optimal in ihrer Entwicklung unterstützt werden sollen, muss die Qualität in diesem Bereich gesteigert werden.

Ein wichtiger Aspekt im Bereich der Lernunterstützung ist die sprachliche Anregung durch die pädagogischen Fachkräfte. Verschiedene Meta-Analysen haben gezeigt, dass das (sprachliche) Interaktionsverhalten von Fachkräften mithilfe von Fort- und Weiterbildungen verbessert werden kann (z.B. Egert, 2015; Markussen-Brown et al., 2017). In der klassischen Lehr-/Lernforschung hat sich Feedback als eine sehr wirksame Methode zur Lernunterstützung herausgestellt (z.B. Hattie & Timperley, 2007; Kluger & DeNisi, 1996). Im frühpädagogischen Kontext wird Feedback zunehmend im Rahmen von Fortbildungsprogrammen eingesetzt. Ein standardisiertes und empirisch begründetes Feedback-Konzept, dessen Effekte wissenschaftlich belegt sind, gibt es bislang im deutschsprachigen Raum allerdings nicht.

Auf der Basis empirischer Hinweise für die Gestaltung von effektivem Feedback wurde deshalb von der Autorin im Rahmen des Projekts „BiSS-Evaluation“<sup>2</sup> ein Konzept (Kappauf, 2018) entwickelt, mithilfe dessen frühpädagogischen Fachkräften ein Feedback zu ihrem sprachlichen Interaktionsverhalten gegeben werden kann.

Ziel der vorliegenden Arbeit war es, die Effekte eines solchen Feedbacks auf das sprachliche Interaktionsverhalten zu überprüfen. Die Evaluation des Feedbacks erfolgte beispielhaft für die sprachliche Anregung während einer Bilderbuchbetrachtung. Da Sprache immer im Kontext von Interaktionen steht und das Feedback-Konzept prinzipiell auch auf andere Verhaltensaspekte übertragen werden kann, liegt zwar der Fokus dieser Arbeit auf der sprachlichen Anregung, jedoch immer eingebettet in das Gesamtkonstrukt Interaktion.

Feedback vergleicht immer einen Ist-Zustand mit einem Soll-Zustand (Müller & Ditton, 2014). Im ersten Theoriekapitel dieser Arbeit wird daher zunächst auf beide Faktoren eingegangen (Kapitel 2.1). Um sich einem Verständnis von guter Interaktionsqualität (Soll-Zustand) zu nähern, wird die Bedeutung von Interaktionen für die kindliche Entwicklung

---

<sup>2</sup> Das Feedback-Konzept wurde von der Autorin im Rahmen des am Staatsinstitut für Frühpädagogik durchgeführten Forschungsprojekts „BiSS-E1 – Alltagsintegrierte Sprachbildung und -diagnostik“ in KITAS: Formative Prozessevaluation der Bund-Länder-Initiative „Bildung durch Sprache und Schrift“ (BIS.00.00001.15) entwickelt, welches im Folgenden zur Vereinfachung als „BiSS-Evaluation“ bezeichnet wird.

betrachtet. Darauf aufbauend werden verschiedene Instrumente zur Operationalisierung und Einschätzung von Interaktionen vorgestellt (Ist-Stand). Anhand von wissenschaftlichen Befunden zur Interaktionsqualität und deren Zusammenhang zur kindlichen Entwicklung werden die theoretischen Annahmen empirisch untermauert und ein Handlungsbedarf aufgezeigt. Anschließend wird auf verschiedene Möglichkeiten eingegangen, wie die Qualität von Interaktionen verbessert werden kann (Kapitel 2.2). Der Fort- und Weiterbildung als wichtige Säule der Qualitätsentwicklung wird dabei besondere Aufmerksamkeit geschenkt.

Im zweiten Teil des theoretischen Hintergrunds wird die Literatur zusammengefasst, auf deren Basis das Feedback-Konzept entwickelt wurde (Kapitel 3). Als Grundlage dient dabei eine begriffliche Bestimmung und Eingrenzung. Auf der Basis theoretischer Modelle und wissenschaftlicher Erkenntnisse zu den Funktionsweisen und Wirkbedingungen von Feedback werden schließlich Implikationen für die Gestaltung von Feedback abgeleitet. Aufbauend darauf ist ein Konzept für ein online-basiertes Videofeedback bezüglich der sprachlichen Anregung von frühpädagogischen Fachkräften entstanden, welches in seinen Grundzügen in Kapitel 5 vorgestellt wird.

Das Ziel der Arbeit bestand darin, die Effekte eines solchen Feedbacks auf das sprachliche Interaktionsverhalten bei einer Bilderbuchbetrachtung zu untersuchen. Wie die Studie methodisch umgesetzt wurde, wird in Kapitel 6 mit Blick auf die erfassten Variablen und die dafür eingesetzten Instrumente, die Stichprobe, den Ablauf der Untersuchung sowie die statistische Auswertung thematisiert. In Kapitel 7 folgt die Beschreibung der Ergebnisse in Hinsicht auf die formulierten Fragestellungen und Hypothesen. Daran schließt sich die Diskussion der Befunde an (Kapitel 8). Abschließend wird in Kapitel 9 ein Fazit aus den gewonnenen Erkenntnissen gezogen und ein zusammenfassender Ausblick für Praxis und Forschung gegeben.



## 2 Pädagogisches Handeln von frühpädagogischen Fachkräften

Als Schlüsselvariable für die Qualität der vorschulischen Bildung, Betreuung und Erziehung von Kindern werden in der wissenschaftlichen Literatur die Interaktionen zwischen den pädagogischen Fachkräften und den Kindern betrachtet (König, 2009; Tietze et al., 1998). Dieses Kapitel beschäftigt sich daher eingehend mit der Interaktionsqualität und Möglichkeiten für deren Optimierung. Dabei wird Interaktionsqualität immer im Allgemeinen und mit speziellem Fokus auf das sprachliche Interaktionsverhalten betrachtet.

### 2.1 Interaktionsqualität als Kernelement von Prozessqualität

Zunächst soll geklärt werden, was gute Interaktionen ausmacht, wie die Qualität gemessen werden kann und in welchem Zusammenhang sie mit der kindlichen Entwicklung steht.

#### 2.1.1 Bedeutung von Interaktionen für die kindliche Entwicklung

Gemäß der „Pädagogik vom Kinde aus“ betonen verschiedene Autoren, dass hinsichtlich der Interaktionsqualität vor allem aus der Perspektive der Kinder gedacht werden sollte (z.B. Katz, 1996). Es wird deshalb in diesem Kapitel zunächst darauf eingegangen, welche Bedeutung Interaktionen für die kindliche Entwicklung haben.

##### Bedeutung von Interaktionen für den Aufbau von Beziehungen

Nach Watzlawick, Bavelas und Jackson (2007) enthalten Interaktionen neben einem inhaltlichen Aspekt immer auch einen Beziehungsaspekt. Der Beziehungsaspekt bestimmt den Auf- und Abbau von Beziehungen. Die täglichen Interaktionen zwischen der Fachkraft und den Kindern beeinflussen demnach, welche Qualität die Beziehung zwischen ihnen hat und sind gleichzeitig Ausdruck dieser. Dem Aufbau von Beziehungen zwischen Kindern und ihren erwachsenen Bezugspersonen hat sich die Bindungstheorie eingehend gewidmet. Unter einer Bindung versteht man dabei eine starke emotionale Beziehung, die nach Bowlby (2010, S. 22) „*ein durch spezifische Faktoren gesteuertes starkes Kontaktbedürfnis gegenüber bestimmten Personen voraus[setzt]*“ und ein stabiles situationsunabhängiges Merkmal darstellt. Bindungsverhalten definiert Bowlby (2010) als jegliches Verhalten, dass auf Nähe zur Bindungsperson ausgerichtet ist und bei inneren Belastungen (z.B. Hunger, Müdigkeit) oder äußeren Stresssituationen (Abwesenheit der Bindungsperson) am deutlichsten gezeigt wird (z.B. indem das Kind durch Weinen auf sich aufmerksam macht oder

sich zur Person hinbewegt). Dieses biologisch festgelegte Bindungsverhaltenssystem sichert das Überleben und die psychische Gesundheit.

Antithetisch zum Bindungsverhalten gibt es den angeborenen Drang, die Umwelt zu erkunden – auch Explorationsverhalten genannt. Nach Grossmann und Grossmann (2014) sind Bindung und Exploration zwei getrennte, aber integrale und sich ergänzende Systeme. Während eine Person die Nähe der Bindungsperson sucht, wenn sie sich unwohl fühlt, bewegt sie sich von dieser erkundend weg, wenn sie sich sicher fühlt (Bowlby, 2002). Eine Bindungsperson dient demnach als sichere Basis, von der aus das Kind die Umwelt erforschen kann, aber auch als sicherer Hafen, zu dem das Kind bei Unsicherheit oder Bedrohung zurückkehren kann.

Entscheidend für den Aufbau von sicheren Bindungen ist das Verhalten der Bezugsperson in den Situationen, in denen das Kind Bindungsverhalten zeigt. Wenn die Signale des Kindes kontinuierlich ignoriert oder falsch interpretiert werden, kann sich keine sichere Bindung entwickeln. Ainsworth, Blehar, Waters und Wall (1978) führen in diesem Zusammenhang das Konzept der Feinfühligkeit ein und meinen damit die Fähigkeit, die kindlichen Signale und Bedürfnisse wahrzunehmen, richtig zu interpretieren sowie prompt und angemessen darauf zu reagieren.

Die Bindungsforschung lieferte in den letzten Jahrzehnten zahlreiche Belege dafür, dass Kinder auch zu ihren pädagogischen Fachkräften eine Bindungsbeziehung aufbauen können (Ahnert, Pinquart & Lamb, 2006). Die Ergebnisse der Meta-Analyse von Ahnert et al. (2006) sprechen dafür, dass das Bindungsmuster einer Beziehung auf den Interaktionserfahrungen mit der jeweiligen Bindungsperson basieren. Genau wie bei den elterlichen Bindungen kommt es demnach auch bei Fachkraft-Kind-Beziehungen auf ein feinfühliges Verhalten der Erzieherin an. Hier muss jedoch der Gruppenkontext in der institutionellen Betreuung berücksichtigt werden. Ein *„empathisches, gruppenbezogen ausgerichtet[es] Erziehverhalten, das die wichtigsten sozialen Bedürfnisse eines Kindes unter der Einbeziehung der Anforderungen der Gruppe zum richtigen Zeitpunkt bedient“*, ist laut Ahnert (2004, S. 268) ausschlaggebend für die Bindungssicherheit eines Kindes zu seiner Erzieherin.

### Bedeutung von Interaktionen für Entwicklungs- und Bildungsprozesse

Basierend auf einer vertrauensvollen Beziehung zur Bezugsperson können Kinder gemäß dem Konzept der sicheren Basis ihre Umwelt erkunden und Entwicklungs- und Lernangebote wahrnehmen und nutzen. Die Rolle von Interaktionen für die kognitive sowie soziale und emotionale Entwicklung soll im Folgenden anhand von bedeutenden Entwicklungstheorien skizziert werden. Da in der vorliegenden Studie im Rahmen des Feedbacks in ers-

ter Linie das sprachförderliche Verhalten von Erzieherinnen reflektiert wird, wird auf die sprachliche Entwicklung verstärkt eingegangen.

Die Entwicklung kognitiver Kompetenzen umfasst den Erwerb von einer Bandbreite an Fähigkeiten, wie z.B. Wahrnehmung, Aufmerksamkeit, logisches Denken und Sprache. Eine wichtige Theorie der kognitiven Entwicklung stammt von Jean Piaget. In der Lerntheorie von Piaget (1992) wird Lernen als Konstruktionsprozess eines selbständigen Individuums gesehen. Piaget betont die aktive Rolle des Kindes bei seiner Entwicklung. Da isolierte Individuen für ihn jedoch nicht existieren, berücksichtigt er gleichzeitig die Bedeutung von anderen Personen für die kindliche Entwicklung. Zwei wichtige Prozesse stehen im Mittelpunkt seiner Theorie: Assimilation und Akkomodation. Eine lernende Person konstruiert sich zunächst immer auf der Basis eingehender Informationen eine eigene Wirklichkeit (Assimilation). Die vorhandenen Wissensstrukturen können dann aufgrund neuer Erfahrungen mit der Umwelt angepasst werden (Akkomodation).

Wygotski (1987) betont noch stärker die Rolle der Umwelt in der kindlichen Entwicklung. Seine Annahmen bildeten den Ursprung sozio-kultureller Theorien, welche die kognitive Entwicklung nicht als isolierten, individuellen Prozess, sondern als das Ergebnis sozialer Interaktion verstehen. In der Interaktion mit anderen Menschen sieht Wygotski (1987, S. 85) die *„allererste Quelle für die Entwicklung der inneren individuellen Eigenschaften der Persönlichkeit des Kindes“*. Die Grundlage der kognitiven Entwicklung wird von Vertretern der sozio-kulturellen Theorien in der Intersubjektivität gesehen. Damit ist das wechselseitige Verständnis in der Kommunikation gemeint. Im Zentrum der Intersubjektivität steht die Fähigkeit zur *„geteilten Aufmerksamkeit“*, welche den Prozess beschreibt, bei dem Interaktionspartner bewusst ihre Aufmerksamkeit auf den gleichen Gegenstand richten (Siegler et al., 2008). Eine hohe Bedeutung für die kognitive Entwicklung von Kindern kommt zudem der sogenannten *„Zone der nächsten Entwicklung“* zu (Wygotski, 1969). Diese bezeichnet den Bereich der Leistung, den das Kind mit einer optimalen Unterstützung erreichen kann. Aufgabe des Erwachsenen ist es, das kindliche Lernen durch Hilfestellungen zu unterstützen, die das Kind weder unter- noch überfordern. Ziel ist es, das Kind zu befähigen, ein Problem selbst zu lösen und damit den nächsten Entwicklungsschritt zu erreichen. Dieses Prinzip findet sich auch im Begriff des *„scaffolding“* wieder (Textor, 2007). Auf den theoretischen Annahmen Wygotskis basiert das Konzept der Ko-Konstruktion, welches in einigen Bildungsplänen Berücksichtigung findet (Fthenakis, 2009). Ko-Konstruktion meint, dass *„Lernen im Rahmen kooperativer und kommunikativer Alltagshandlungen und Bildungsaktivitäten geschieht, an denen Kinder und Erwachsene gleichermaßen aktiv beteiligt sind“* (Reichert-Garschhammer & Kieferle, 2011, S. 88). Wissen wird in diesem Sinne gemeinsam konstruiert, weshalb Wissenserwerb und Kompetenzentwicklung das Ergebnis sozialer Pro-

zesse sind. Die Interaktion zwischen Kind und Erzieherin ist demnach der Schlüssel für Ko-Konstruktion. Gemäß der Definition von Ko-Konstruktion heben Siraj-Blatchford (2007) sowie Pramling Samuelsson (2009) hervor, dass eine aktive Beteiligung der erwachsenen Person eine grundlegende Voraussetzung für Lernprozesse ist. Eine aktive Beteiligung und Anregung im Sinne der Ko-Konstruktion meint jedoch nicht, dass die Fachkraft den Interaktionsablauf dominiert und Kinder instruiert. Nach einem sozial-konstruktivistischen Bildungsverständnis orientiert sich die Qualität der Interaktionen einerseits daran, wie aktiv und selbstbestimmt sich das Kind mit seiner Lernumgebung auseinandersetzen kann und andererseits daran, inwiefern die pädagogische Fachkraft das Kind bei diesen Lernprozessen unterstützt und anregt. Dies entspricht auch der Theorie von Deci und Ryan (1993), nach der Selbstbestimmung eines von drei psychologischen Grundbedürfnissen ist. Ein weiterer entwicklungsförderlicher Aspekt von Interaktionen, welcher im Zusammenhang der ko-konstruktiven Interaktion häufig genannt wird, ist das „sustained shared thinking“. Sylva, Melhuish, Sammons, Blatchford und Taggart (2014, S. 6) definieren „sustained shared thinking“ wie folgt: *„Sustained shared thinking is where two or more individuals work together in an intellectual way to solve a problem, clarify a concept, evaluate an activity, extend a narrative etc. Both parties must contribute to the thinking and it must develop and extend the understanding“*.

Das Erlernen von Sprache stellt einen wichtigen Teil der kindlichen Entwicklung dar. Unter sprachlicher Kompetenz wird sowohl das Sprachverstehen als auch die Sprachproduktion subsumiert. Der Spracherwerb umfasst das Erlernen vieler verschiedener Komponenten von Sprache. Im Großen und Ganzen zählen dazu die phonologische (Lautsystem der Sprache), die semantische, die syntaktische und die pragmatische Entwicklung (Wissen darüber, wie Sprache verwendet wird). Es gibt einige empirische Hinweise darauf, dass die frühen Lebensjahre eine kritische Phase in der sprachlichen Entwicklung darstellen. Demnach erlernen Kinder, vor allem in den ersten Jahren, deutlich leichter Sprache als Erwachsene. Neben den neuronalen Voraussetzungen im Gehirn, spielt die soziale Umwelt eine entscheidende Rolle bei der Sprachentwicklung. Kinder müssen Kontakt zu Menschen haben, die Sprache verwenden, um selbst Sprache zu erlernen (Milburn, Girolametto, Weitzman & Greenberg, 2014; Siegler et al., 2008). Spracherwerb ist daher ein Musterbeispiel für Ko-Konstruktion (Reichert-Garschhammer & Kieferle, 2011). Wichtige Bezugspersonen, wie die Eltern oder das Betreuungspersonal in der Kita, können somit die sprachliche Entwicklung von Kindern aktiv unterstützen. Damit dies gelingen kann, müssen bestimmte Voraussetzungen in der Kommunikation mit dem Kind erfüllt sein. Die geteilte Aufmerksamkeit zwischen Erwachsenem und Kind ist unverzichtbar für den Erwerb von Sprache (Tomasello, 2009). Zwei weitere wichtige Aspekte für das Erlernen von Sprache sind Blickkontakt und

körperliche Zuwendung. Zum einen kann über die körperliche Zugewandtheit und das Herstellen von Blickkontakt ein gemeinsamer Aufmerksamkeitsfokus hergestellt werden. Zum anderen kann das Kind so Mimik und Gestik seines Gegenübers beobachten, welche ihm dabei helfen, die sprachlichen Inhalte zu interpretieren (Reichert-Garschhammer & Kieferle, 2011). Kinder erlernen Sprache vor allem durch die aktive Verwendung in Handlungen und Dialogen mit Interaktionspartnern (Bruner, 2008). Im Sinne sozio-kultureller Theorien ist Sprache im Allgemeinen der Schlüssel zum Lernen, da in der verbalen Kommunikation ko-konstruktiv Bedeutungen erschlossen werden (Hetherington, Parke & Schmuckler, 2005). In diesem Zusammenhang ist es wichtig, Gesprächsanlässe im Alltag zu erkennen und zu nutzen. Eine Möglichkeit, Kinder aktiv zum Sprechen anzuregen, ist das Stellen von Fragen. In der Regel werden geschlossene und offene Fragen unterschieden. Die Definitionen variieren zum Teil sehr stark und berücksichtigen dabei unterschiedliche Aspekte. Zum einen kann danach differenziert werden, wie weit oder eng die Frageformen gefasst werden. In einem engen Verständnis zählen nur sogenannte „Ja-/Nein-Fragen“ zu den geschlossenen Fragen, also Fragen, auf die das Kind nur mit „Ja“ oder „Nein“ antworten kann (Schröder & Keller, 2013). Das weitere Verständnis bezieht alle Fragen mit ein, auf die nur mit einem Wort oder einer kurzen Äußerung geantwortet werden kann (z.B. Wasik, Bond & Hindman, 2006; Kucharz, Mackowiak & Beckerle, 2015), also auch „Alternativfragen“ (z.B. „Ist das ein Hund oder eine Katze?“), rhetorische Fragen (z.B. „Hab ich es euch nicht gesagt?“) und Faktenfragen (z.B. „Was ist das?“). Offene Fragen sind nach diesem Verständnis hingegen so gestellt, dass das Kind zu längeren Antworten angeregt wird (z.B. „Was passiert auf dieser Seite?“). In einzelnen Studien konnte gezeigt werden, dass offene Fragen von Fachkräften tatsächlich zu längeren Äußerungen bei Kindern führen als geschlossene Fragen (z.B. De Rivera, Girolametto, Greenberg & Weitzman, 2005). Andere Wissenschaftler beziehen sich in ihrer Definition auf die inhaltlichen Antwortmöglichkeiten (z.B. Milburn et al., 2014). Danach haben geschlossene Fragen nur sehr eingeschränkte Antwortmöglichkeiten, wie z.B. Ja/Nein-Fragen oder Alternativfragen. Dazu zählen beispielsweise auch Fragen, bei denen die Antwort bekannt ist und welche demnach nur als Abfrage dienen. Im Gegensatz dazu ist bei offenen Fragen die Antwort inhaltlich nicht eingeschränkt oder vorbestimmt. Offenen Fragen wird in diesem Zusammenhang auch zugeschrieben, dass sie Kinder stärker kognitiv anregen (z.B. „Warum ist der Junge traurig?“) und so ein „sustained shared thinking“ ermöglichen können (Massey, Pence, Justice & Bowles, 2008). Siraj-Blatchford und Manni (2008) integrieren beide Ansätze und berücksichtigen sowohl die Länge als auch die inhaltliche Einschränkung der möglichen Antwort.

Um die sprachlichen Kompetenzen zu erweitern, muss das Kind innerhalb der „Zone der nächsten Entwicklung“ unterstützt werden. So kann bei Kindern mit geringen sprachlichen

Fähigkeiten zum Beispiel eine geschlossene Frage hilfreicher sein als eine offene Frage (Blewitt, Rump, Shealy & Cook, 2009). Ein korrekatives Feedback zu den kindlichen Äußerungen bietet ebenfalls Gelegenheit, Kinder in ihrer „Zone der nächsten Entwicklung“ anzuregen. Wenn ein Kind beispielsweise noch nicht in der Lage ist, die richtigen Artikel zu verwenden (z.B. Kind sagt „Die Baum“), kann die Erzieherin die Äußerung korrigierend wiederholen („Der Baum“). Sprechen Kinder noch in Ein- oder Zwei-Wortäußerungen, kann die Fachkraft die Aussage in einem ganzen Satz wiederholen und auf diese Weise die Sprache des Kindes erweitern. Für die Unterstützung des Sprachverständnisses können Interaktionspartner die eigenen Handlungen oder die der Kinder mit Sprache begleiten. Durch die Verknüpfung der Handlung mit dem passenden Inhalt, fällt es dem Kind leichter, die sprachlichen Inhalte zu verstehen. Dies gilt vor allem für Kinder in den Anfangsstadien des Spracherwerbs (Reichert-Garschhammer & Kieferle, 2011).

Für die Entwicklung von sozialen Fähigkeiten spielen Interaktionen ebenfalls eine herausragende Rolle. Nach der psychoanalytischen Theorie von Sigmund Freud entwickelt sich sozial kompetentes Verhalten aus dem Kompromiss zwischen der eigenen Bedürfnisbefriedigung und der Berücksichtigung der internalisierten Wertvorstellungen der Umwelt (Siegler et al., 2008). Auch die sogenannten Lerntheorien gehen davon aus, dass sich die soziale Entwicklung hauptsächlich auf das Lernen in Interaktionen mit anderen Menschen zurückführen lässt. Dabei kann man die behavioristischen Theorien und die Soziale Lerntheorie unterscheiden. Die behavioristischen Theorien betrachten Konditionierung als den zentralen Lernmechanismus. Kinder erlernen dann sozial kompetentes Verhalten, wenn es vom Interaktionspartner verstärkt wird (Skinner, 1968). Soziale Lerntheorien gehen eher von Beobachtung und Nachahmung und weniger von Verstärkung als Entwicklungsmechanismen aus. Nach Bandura (1979), dem Begründer der sozialen Lerntheorie, lernen Kinder am schnellsten und wirksamsten, indem sie beobachten, wie sich andere Menschen verhalten und dieses Verhalten dann nachahmen. Verstärkung kann dabei die Wahrscheinlichkeit erhöhen, dass das Verhalten tatsächlich nachgeahmt wird.

Die emotionale Entwicklung von Kindern wird vor allem durch die Qualität der frühen sozialen Beziehungen bestimmt. Im Austausch mit engen Bezugspersonen erlangen Kinder eine Repräsentation von Gefühlen in bestimmten Situationen und Wissen über emotionales Ausdrucksverhalten. Ein zentraler Schritt in der emotionalen Entwicklung ist die Entwicklung der Emotionsregulierung. Kleinkinder sind zunächst von der Unterstützung ihrer Bezugspersonen abhängig, um emotionale Erregungszustände regulieren zu können. Durch ein einfühlsames Wiegen oder Trösten kann die emotionale Erregung von Kindern verringert werden. Man spricht dabei von einer „Ko-Regulation“. In diesem Zusammenhang spielt das feinfühlige Verhalten der Bezugsperson eine wesentliche Rolle. Je feinfühligere die Be-

zugsperson auf die emotionale Erregung des Kindes reagiert, desto besser ist das Kind in der Lage, mit der Zeit eigene Regulationsstrategien zu entwickeln (Weltzien, 2014).

Verschiedene theoretische Ansätze betonen, dass neben genetischen Voraussetzungen die soziale Umwelt eine zentrale Rolle im Rahmen der kindlichen Entwicklung spielt. Kinder bauen über Interaktionen eine Beziehung zur pädagogischen Fachkraft auf, welche die Funktion einer Bindungsbeziehung erfüllen kann. In der Interaktion mit der Erzieherin haben Kinder die Möglichkeit, wichtige kognitive und sozial-emotionale Kompetenzen zu erlernen. Entscheidend ist daher, wie pädagogische Fachkräfte diese Prozesse unterstützen. Was ein entwicklungsförderliches Interaktionsverhalten ausmacht, wurde in den vergangenen Abschnitten bereits angerissen. Neben der Feinfühligkeit von Erzieherinnen, wurden beispielsweise die „Zone der nächsten Entwicklung“ und das damit verbundene „scaffolding“-Konzept sowie verschiedene sprachanregende Strategien genannt.

## 2.1.2 Operationalisierung von Fachkraft-Kind-Interaktionen

Zahlreiche Autoren haben versucht, die Indikatoren für positive Interaktionen zu operationalisieren, um Interaktionsqualität messen zu können. In der frühpädagogischen Forschung ist daher über die Zeit hinweg eine Vielzahl an verschiedenen Beobachtungs- und Einschätzverfahren entstanden, mit denen die Interaktionsqualität zwischen Fachkräften und Kindern beurteilt werden kann. Auf die unterschiedlichen Ansätze soll in diesem Kapitel eingegangen werden.

### Operationalisierung im Rahmen globaler Prozessqualität

Wie bereits deutlich wurde, sind die Interaktionen zwischen pädagogischen Fachkräften und Kindern ein Teil der Prozessqualität in der Kita. Zur Erfassung der Prozessqualität wurde international bisher vor allem die „Early Childhood Environment Rating Scale“ (ECERS-R) von Harms, Clifford und Cryer (1998) eingesetzt. Die ECERS-R bewertet sieben Qualitätsdimensionen: Platz und Ausstattung, Betreuung und Pflege der Kinder, sprachliche und kognitive Anregungen, Aktivitäten, Interaktionen, Strukturierung der pädagogischen Arbeit sowie Eltern und Erzieherinnen. Die Einschätzung erfolgt auf einer siebenstufigen Skala. Im Rahmen der EPPE-Studie wurde eine Erweiterung der ECERS-R entwickelt: die ECERS-E (Sylva, Siraj-Blatchford & Taggart, 2011). Die ECERS-E bezieht zusätzliche Merkmale aus den vier Bereichen Literacy und Lesen, Mathematik, Naturwissenschaften und Umwelt sowie Vielfalt und individuelle Förderung ein. Gemeinsam mit seinen Mitarbeitern übersetzte Tietze die ECERS-R ins Deutsche, wobei der Inhalt im Wesentlichen nicht verändert wurde. Die deutsche Adaption der ECERS-R ist unter dem Namen „Kindergarten-Einschätz-Skala“

(KES-R; Tietze, Schuster, Grenner & Roßbach, 2001) bekannt. Durch die eher globale Perspektive auf das Phänomen der Prozessqualität ist eine differenzierte Analyse der spezifischen Struktur der Interaktionen zwischen Fachkraft und Kind bei diesen Instrumenten nicht möglich. Zudem kritisieren einige Autoren (z.B. Gordon, Fujimoto, Kaestner, Korenman & Abner, 2013; Mayer & Beckh, 2016) die ECERS-R und die KES-R hinsichtlich ihrer schwachen psychometrischen Eigenschaften und der niedrigen Validität zur Vorhersage von kindlicher Entwicklung.

### Operationalisierung mit dem Fokus auf Beziehungs- und Bindungsaspekte

Die „Fremde Situation“, welche ursprünglich zur Einschätzung der Qualität der Mutter-Kind-Bindung entwickelt wurde, wurde in der Forschung auch auf Erzieherinnen-Kind-Beziehungen übertragen. Das Bindungsverhalten des Kindes nach der Trennung von der Erzieherin gibt dabei analog Auskunft über die Bindungssicherheit. Dabei werden vier Bindungsmuster unterschieden: sicher, unsicher-vermeidend, unsicher-ambivalent, desorganisiert-desorientiert (Ainsworth et al., 1978; Main & Solomon, 1986). Ahnert (2004) weist in diesem Zusammenhang darauf hin, dass das Fehlen einer Bindungsbeziehung zwischen Kind und Erzieherin in der „Fremden Situation“ als unsicher-vermeidende Bindung fehlinterpretiert werden kann. So muss vermeidendes Verhalten des Kindes nicht unbedingt auf eine unsicher-vermeidende Bindung hinweisen, sondern könnte auch schlichtweg bedeuten, dass die Beziehung keine Bindungsfunktion erfüllt und es sich eher um Vermeiden im Sinne von Desinteresse handelt.

Neben der „Fremden Situation“ wird auch das „Attachment Q-Sort“ (AQS) von Waters und Deane (1985) häufig eingesetzt, um die Bindungssicherheit der Beziehung von Kindern zu einer Betreuungsperson einzuschätzen. Da das AQS ursprünglich zur Messung der Bindung von jungen Kindern zu ihren Müttern entwickelt wurde, wurde es für Beziehungen zu anderen Betreuungspersonen adaptiert und konnte so auch zu deren Messung eingesetzt werden (z.B. Howes, Matheson & Hamilton, 1994). Bei der „Q-Sort-Methode“ wird beobachtet, inwiefern Kinder die Bindungsperson als „sichere Basis“ benutzen (Waters & Deane, 1985). In der überarbeiteten Version von Waters (1995) müssen die Beobachter dabei 90 Aussagen, welche das Verhalten der Kinder während der Interaktion mit einer Betreuungsperson beschreiben, auf einer neunstufigen Skala beurteilen. Die Items beschreiben fünf verschiedene Aspekte einer Bindungsbeziehung: Sicherheit, Assistenz, Zuwendung, Stressreduktion und Explorationsunterstützung. Sowohl die „Fremde Situation“ als auch das AQS wurden vielfach erfolgreich zur Messung der Bindungssicherheit zwischen Erzieherin und Kind eingesetzt und erfüllten dabei grundlegende statistische Gütekriterien (Ahnert et al., 2006). Während die „Fremde Situation“ hauptsächlich ihren Fokus auf die Trennungsbelastung des Kindes legt, deckt das AQS ein breiteres Spektrum an Bindungsaspekten ab und ist



daher im Kontext von außerfamiliären Betreuungskontexten besser geeignet (Ahnert, 2004; Ahnert, 2010).

Basierend auf bindungstheoretischen Annahmen entwickelte Pianta (2001) eine Ratingskala, mit der Erzieher und Lehrer die Qualität der Beziehung zu einem bestimmten Kind selbstevaluativ einschätzen können: die „Student-Teacher-Relationship Scale“ (STRS). Das Beobachtungsinstrument eignet sich für Kinder zwischen ca. vier bis neun Jahren. Die bis heute bestehende Fassung der STRS setzt sich aus 28 Items zusammen und untersucht die Dimensionen „conflict“, „closeness“ und „dependency“. Die STRS wurde bis heute in zahlreichen Untersuchungen eingesetzt, in denen ihre guten psychometrischen Eigenschaften immer wieder bestätigt wurden (z.B. Howes et al., 2008; Pianta & Stuhlman, 2004).

### Operationalisierung mit dem Fokus auf das Verhalten der Fachkräfte

Verschiedene Ansätze und Verfahren fokussieren noch stärker das Verhalten der Erzieherin in der Interaktion mit den Kindern. Einige davon sollen im Folgenden vorgestellt werden.

Als wesentliche Voraussetzung für eine sichere Bindung gilt ein feinfühliges Verhalten der Bezugsperson. Ahnert (2004) unterscheidet zwischen kindzentriertem und gruppenorientiertem Erzieherinnenverhalten. Zur Beobachtung eines kindzentrierten Verhaltens wird häufig die klassische Sensitivitätsskala von Ainsworth, Bell und Stayton (1974) eingesetzt. Die Skala wurde dafür entwickelt, das sensitive Verhalten der Mutter während der Interaktion mit dem Kind auf einer neun-stufigen Skala (1 = „highly insensitive“ bis 9 = „highly sensitive“) einzuschätzen. Howes und Stewart (1987) beschreiben mit ihrer „Adult Involvement Scale“ auf sieben Stufen, wie responsiv die Erzieherin in das Gruppengeschehen involviert ist. Die unterste Stufe („ignore“) beschreibt eine fehlende Beteiligung und Verfügbarkeit der Fachkraft (z.B. durch Abwesenheit oder organisatorische Aufgaben). Erzieherinnen auf der höchsten Stufe („intense“) sind physisch und verbal verfügbar für einen Großteil der Gruppe, gehen auf kindliche Handlungen oder Äußerungen ein und initiieren Gespräche. Gruppenorientierte Verfahren scheinen dabei besser dazu geeignet zu sein, die Bindungssicherheit von Kindern zu pädagogischen Fachkräften vorherzusagen (Ahnert, 2004). Remsperger (2011) beschreibt ein responsives Erzieherinnenverhalten anhand folgender Merkmale: Zugänglichkeit, Aufmerksamkeit, Promptheit der Reaktion, Richtigkeit der Interpretation, Generelle Haltung (Akzeptanz, Interesse und Respekt vor Autonomie), Involvement, Emotionales Klima und Stimulation und erweitert damit das klassische Konzept von Feinfühligkeit.

Das mit am häufigsten eingesetzte Instrument zur Erfassung der Qualität von Interaktionsverhalten ist die „Caregiver Interaction Scale“ (CIS) von Arnett (1989). Die CIS besteht aus 26 Items, welche sich auf vier Dimensionen verteilen: „Positive Interaction“, „Punitiveness“, „Detachment“ und „Permissiveness“. Der Faktor „Positive Interaction“ umfasst Aspekte wie

die Wärme der Interaktion oder die entwicklungsangemessene Kommunikation mit dem Kind. Unter „Punitiveness“ versteht Arnett ablehnendes, drohendes oder sehr kritisches Verhalten gegenüber dem Kind. Der Faktor „Detachment“ erfasst, in welchem Ausmaß die Erzieherin kein Interesse am Kind zeigt, also alle Aktivitäten, die keine Interaktion mit dem Kind enthalten. Der Faktor „Permissiveness“ definiert schließlich die Toleranz der Erzieherin in Bezug auf das Fehlverhalten des Kindes (Arnett, 1989). Tietze et al. (1998) haben in ihrer Studie eine ins Deutsche übersetzte Form der „Caregiver Interaction Scale“ verwendet. Im Zuge einer Faktorenanalyse der deutschen CIS konnte der Faktor „Permissiveness“ nicht reproduziert werden, was zu einer drei-faktoriellen Struktur mit den Subskalen „Sensitivität“, „Involviertheit/Beteiligung“ und „Akzeptanz“ führte (Tietze et al., 1998).

Die Forschungsgruppe um Pianta entwickelte ein konzeptuelles Rahmenmodell von Interaktionsqualität. Das TTI-Modell („Teaching through Interactions“) unterscheidet drei große Bereiche von Fachkraft-Kind-Interaktionen: Emotionale Unterstützung („Emotional support“), Organisation des Kita-Alltags („Classroom Organization“) und Lernunterstützung („Instructional support“) (Hamre et al., 2013). Diese Dimensionen finden sich auch in dem dazugehörigen Beobachtungsinstrument, dem „Classroom Assessment Scoring System“ (CLASS Pre-K; Pianta, La Paro & Hamre, 2008) wieder. Konfirmatorische Faktorenanalysen, welche auf den Daten von über 4000 Kindergartengruppen und Schulklassen beruhen, weisen darauf hin, dass die vorhergesagte drei-faktorielle Struktur die Interaktionsqualität am besten abbildet. In Abbildung 1 sind die drei Faktoren mit ihren jeweiligen Dimensionen dargestellt.

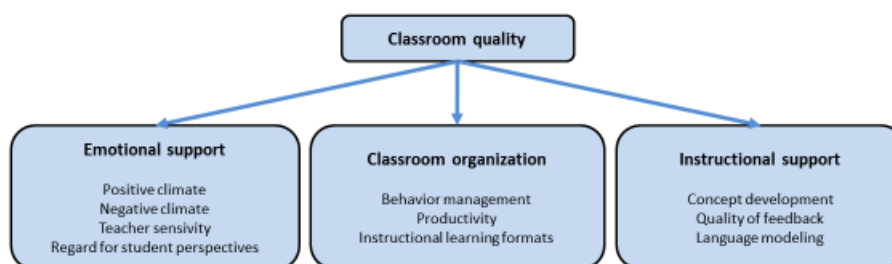


Abbildung 1: Faktoren und Dimensionen des "Classroom Assessment Scoring System" (CLASS Pre-K; Pianta, La Paro et al., 2008)

Die einzelnen Dimensionen werden wiederum anhand verschiedener Indikatoren und ihrer jeweiligen „Behavioral Marker“ beschrieben, mithilfe derer man die Ausprägung der Dimension einschätzt. Jede der 10 Dimensionen wird auf einer sieben-stufigen Ratingskala bewertet. Die Skala wird in drei Qualitätsbereiche unterteilt („Low“ = 1, 2; „Mid“ = 3, 4 und 5; „High“ = 6, 7). Die Einschätzung erfolgt durch Live-Beobachtungen mit mehreren „obser-

vation cycles“. Jeder „cycle“ dauert ca. 20 Minuten. Pro „observation cycle“ wird ein Wert für jede Dimension vergeben. Die CLASS wurde dafür entwickelt, die Interaktionsqualität in einem Gruppensetting zu beobachten. Es geht um die durchschnittliche Erfahrung eines Kindes in dieser Gruppe. Die Autoren verweisen jedoch darauf, dass das Instrument auch dafür verwendet werden kann, um speziell das Interaktionsverhalten einer pädagogischen Fachkraft zu beobachten und dass dieses auch anhand von Videoaufnahmen anstatt von Live-Beobachtungen geschehen kann (Pianta, La Paro et al., 2008). Die Reliabilität und Validität der CLASS wurde in verschiedenen großen Studien getestet. Pianta und La Paro et al. (2008) geben eine durchschnittliche Übereinstimmung von 87 % zwischen zwei Ratern an. Eine Übereinstimmung wird angenommen, wenn die Werte  $\pm 1$  auseinanderliegen. Verschiedene Untersuchungen ergaben signifikante Zusammenhänge zu anderen Qualitätsmaßen (z.B. ECERS-R) und Aspekten der kindlichen Entwicklung (z.B. Howes et al., 2008; Mashburn et al., 2008). Bisherige Erfahrungen mit dem Einsatz der CLASS in Europa (z.B. Pakarinen et al., 2010; Wildgruber, Wirts & Wertfein, 2014) zeigen, dass ihre Dimensionen auch auf die Interaktionsqualität in europäischen bzw. deutschen Kindertageseinrichtungen übertragbar sind.

Die meisten Instrumente zur Erfassung der Interaktionsqualität wurden ursprünglich für den englischsprachigen Raum entwickelt. Basierend auf einem theoriegeleiteten Diskurs sowie der Analyse von Videosequenzen aus dem Alltag von pädagogischen Fachkräften in Deutschland entwickelte Weltzien (2014) den Beobachtungs- und Reflexionsbogen „GlnA“ (Gestaltung von Interaktionsgelegenheiten im Alltag). Der Bogen besteht aus drei Skalen mit insgesamt 28 Merkmalen: „Gesprächsbereitschaft“ (7 Merkmale), „Gestaltung der Interaktion“ (10 Merkmale) und „Aktivierung von Fachwissen“ (11 Merkmale). Jedes Merkmal wird mithilfe einer vier-stufigen Skala eingeschätzt, wobei immer die zwei gegensätzlichen Ausprägungen inhaltlich beschrieben sind. Auf der Grundlage von 60 Videosequenzanalysen wurde der Beobachtungsbogen auf seine Skaleneigenschaften überprüft. Die faktorenanalytische Überprüfung und die Reliabilitätsanalysen bestätigten die drei-faktorielle Struktur. Für einige Items stellte sich eine zu geringe Item-Schwierigkeit heraus, diese wurden jedoch aus inhaltlichen Gründen beibehalten. Die Inter-Rater-Reliabilität wurde anhand von 28 Videosequenzen analysiert, welche jeweils durch zwei Rater bewertet wurden. Es ergab sich eine mäßige bis gute Übereinstimmung, was von der Autorin als noch nicht zufriedenstellend betrachtet wird (Weltzien, 2013).

Ein Verfahren, welches vor allem sprachanregende Interaktionen von Erzieherinnen in den Blick nimmt, ist die „Dortmunder Einschätzskala zur Erfassung sprachförderrelevanter Interaktionen“ (DO-RESI; Fried & Briedigkeit, 2008). Die DO-RESI besteht aus vier Dimensionen mit insgesamt 23 Items: „Organisation der Bildungsangebote“ (z.B. Planen/Routinen),

„Emotionale Beziehung“ (z.B. Nähe und Empathisches Zuhören), „Adaptive Unterstützung“ (z.B. Sensitivität, Engagiertheit) und „Sprachlich-kognitive Herausforderung“ (z.B. vielfältiger Wortschatz und offene Fragen). Jedes Item wird auf einer sieben-stufigen Skala bewertet (1 = ungenügend bis 7 = exzellent). In einzelnen Studien wurde die Messgüte des Verfahrens in Form von Reliabilität und faktorieller bzw. kriteriumsorientierter Validität (Korrelation mit anderen Instrumenten, wie ECERS-R und CIS) bestätigt (z.B. Fried, 2011).

Die „Teacher Interaction and Language Rating Scale“ (TILRS) von Girolametto, Weitzmann und Greenberg (2000) ist ein englischsprachiges Instrument zur Einschätzung des Trainingsbedarfs von pädagogischen Fachkräften hinsichtlich ihrer sprachförderlichen Interaktionen. Die Verhaltensweisen (z.B. Fragen stellen oder Erweitern der kindlichen Äußerungen) werden auf einer sieben-stufigen Skala eingeschätzt. Werte zwischen „1“ und „3“ bedeuten, dass das Verhalten fast gar nicht bzw. teilweise gezeigt wird und eine Verbesserung dringend nötig ist. Wird eine „4“ vergeben, ist die Veränderung von Feinheiten notwendig. Werte im oberen Bereich (5 bis 7) zeigen an, dass das Verhalten zufriedenstellend umgesetzt wird und nicht während des Trainings thematisiert werden muss.

Über die genannten Beobachtungsverfahren hinaus gibt es noch zahlreiche andere Instrumente, die an dieser Stelle nicht vorgestellt werden können (z.B. „Caregiver Interaction Profile Scales“ von Helmerhorst, Riksen-Walraven, Vermeer, Fukkink & Tavecchio, 2014 oder „Early Language and Literacy Classroom Observation“ von Smith & Dickinson, 2002).

### **2.1.3 Empirische Befunde zur Interaktionsqualität und ihrem Zusammenhang zur kindlichen Entwicklung**

Im vorangegangenen Abschnitt wurde anhand verschiedener Beobachtungs- und Einschätzverfahren aufgezeigt, wie Interaktionsqualität operationalisiert werden kann. Diese und andere Instrumente wurden in zahlreichen Studien eingesetzt, um die Interaktionsqualität in Kindertageseinrichtungen zu erfassen sowie potenzielle Einflussfaktoren und Zusammenhänge zur kindlichen Entwicklung aufzudecken. Die bestehenden wissenschaftlichen Befunde sollen anhand ausgewählter Studien dargelegt werden.

#### Zusammenhang zur kindlichen Entwicklung

Bereits in Kapitel 2.1.1 wurde deutlich, welche Rolle Interaktionen für die kindliche Entwicklung spielen. Mittlerweile wurde in einer Vielzahl an Studien untersucht, in welchem Zusammenhang die Betreuungsqualität allgemein und speziell Fachkraft-Kind-Interaktionen zur kindlichen Entwicklung stehen und welche Aspekte des Interaktionsverhaltens von Erzieherinnen besonders förderlich für die Entwicklung der Kinder sind.

### *Prozessqualität*

In groß angelegten Untersuchungen zur Prozessqualität in Kindertageseinrichtungen wurde die Qualität außerfamiliärer Betreuung in Verbindung mit kindlichen Entwicklungsmaßen gebracht. In der sogenannten NICHD-Studie (NICHD Early Child Care Research Network, 2002) wurde die kognitive und soziale Entwicklung von über 1000 Kindern in den USA längsschnittlich beobachtet. Es zeigte sich, dass die außerfamiliären Betreuungserfahrungen zwar einen deutlich geringeren Einfluss auf die Entwicklung hatten als familiäre Aspekte, aber dennoch einen bedeutenden Prädiktor darstellten. Kinder, welche eine Einrichtung mit hoher Qualität besuchten, zeigten eine höhere kognitive Leistung und weniger Problemverhalten. Ähnliche Ergebnisse ergab auch die Längsschnitt-Studie „Effective Provision of Preschool Education“ (EPPE; Sylva, 2017) in England. Die Wissenschaftler begleiteten 2800 Kinder aus 141 Kindertageseinrichtungen ab dem Alter von drei Jahren. Die Prozessqualität in den Einrichtungen wurde mit der ECERS-R erfasst. Eine höhere Qualität in der Kindertagesbetreuung war mit höheren kognitiven Kompetenzen und einem besseren Sozialverhalten im Alter von fünf, sieben und elf Jahren verbunden. In der NUBBEK-Studie („National Study of Child Care in Early Childhood“) wurde die deutsche Version der ECERS-R, die KES-R, zur Vorhersage kindlicher Entwicklungsmaße verwendet. Insgesamt konnten zwar signifikante Zusammenhänge zur kindlichen Entwicklung gefunden werden, diese waren jedoch nur von geringer Stärke, was nach Mayer und Beckh (2016) unter anderem auf unzureichende psychometrische Eigenschaften der ECERS-E zurückzuführen ist (z.B. geringe strukturelle Validität aufgrund von verschiedenen inhaltlichen Aspekten in einem Item).

### *Bindungs- und Beziehungsqualität*

Howes, Hamilton und Matheson (1994) erforschten die Verbindung zwischen Bindungssicherheit und Sozialkompetenz. Die Analysen ergaben, dass die mit dem AQS erfasste Bindungsqualität der ein- bis dreijährigen Kinder zu ihren außerfamiliären Betreuungspersonen ein signifikanter Prädiktor für ihr soziales Verhalten gegenüber Peers und Erwachsenen im Alter von vier Jahren war. Glüer (2013) untersuchte den Einfluss der Bindung zur Erzieherin auf die Bildungsbereitschaft. In seiner Studie mit 75 Kindern im Kindergarten konnte die Bildungsbereitschaft von Kindern durch die Bindungssicherheit vorhergesagt werden, welche ebenfalls mit dem AQS gemessen wurde.

Als Teil der groß angelegten amerikanischen „Cost, Quality, and Child Outcomes (CQO) in Child Care Centers Study“ haben Peisner-Feinberg et al. (2001) längsschnittlich die kognitive und sozial-emotionale Entwicklung von 733 Kindern im Alter von vier bis acht Jahren in Abhängigkeit ihrer Erfahrungen in der außerfamiliären Betreuungseinrichtung bzw. Schule beobachtet. Zur Einschätzung der Qualität der Erzieherin-Kind-Beziehung nutzten sie die

Skala „Closeness“ aus der STRS. Die kognitiven und sozialen Fähigkeiten standen zu allen Messzeitpunkten in einem signifikanten Zusammenhang zur Nähe in der Erzieherin-Kind-Beziehung bzw. konnten durch sie vorhergesagt werden. Howes et al. (2008) konnten zeigen, dass die emotionale Nähe innerhalb der Beziehung von Erzieherin und Kind in der Vorschulzeit signifikant positiv mit der Lesefertigkeit der Kinder in Verbindung steht. In zahlreichen weiteren Studien (z.B. Ewing & Taylor, 2009; Glüer, 2013; Burchinal, Peisner-Feinberg, Pianta & Howes, 2002; Pianta & Stuhlman, 2004) konnte bestätigt werden, dass die Qualität der Erzieherin-Kind-Beziehung ein wichtiger Prädiktor für die kognitiven bzw. akademischen Kompetenzen von Kindern ist.

### *Interaktionsverhalten*

Auch Studien, welche speziell das Interaktionsverhalten der Fachkräfte in den Blick genommen haben, konnten signifikante Zusammenhänge zur kindlichen Entwicklung feststellen. So konnte beispielsweise eine hohe Qualität der mit der CLASS gemessenen Interaktionsqualität in Verbindung mit einer positiven Entwicklung gebracht werden. Mashburn et al. (2008) untersuchten an einer Stichprobe von 2439 Kindern aus 671 Kindergartengruppen die Qualität der Interaktionen sowie akademische, sprachliche und soziale Kompetenzen am Anfang und am Ende eines Kindergartenjahres. Positive Interaktionen aus dem Bereich „instructional support“ konnten die kognitiven und sprachlichen Fähigkeiten der Kinder signifikant vorhersagen. Ein hohes Ausmaß an „emotional support“ führte zu einer höheren Sozialkompetenz und weniger Problemverhalten. In einer Studie von Rimm-Kaufman, Curby, Grimm, Nathanson und Brock (2009) wurde ein Zusammenhang zwischen allen drei CLASS-Dimensionen und dem adaptiven Verhalten von Kindern beobachtet. Besonders konsistente Ergebnisse gab es für die Dimension „Classroom Organization“. Wurden Kinder von Fachkräften mit einem guten Gruppenmanagement betreut, äußerte sich dies in einem besseren adaptiven Gruppenverhalten der Kinder, z.B. in Form einer besseren Selbstkontrolle und einer höheren Engagiertheit bei Gruppenaktivitäten. Ähnliche Ergebnisse zum Zusammenhang der CLASS-Dimensionen mit kindlichen Entwicklungsmaßen finden sich auch in anderen Studien (z.B. Bailey, Denham, Curby & Bassett, 2016; Broekhuizen, Mokrova, Burchinal & Garrett-Peters, 2016; Curby, Rimm-Kaufman & Ponitz, 2009).

Einige Verhaltensweisen von Erzieherinnen haben sich in wissenschaftlichen Untersuchungen als besonders effektiv für die Unterstützung der kindlichen Entwicklung erwiesen. Ein feinfühliges Verhalten pädagogischer Fachkräfte konnte beispielsweise mit besseren sozialen und kognitiven Kompetenzen von Kindern in Verbindung gebracht werden. Demnach führt ein feinfühliges Interaktionsverhalten zu weniger Verhaltensproblemen (NICHD Early Child Care Research Network, 2003; Rimm-Kaufman et al., 2002) und einer höheren Sozialkompetenz bei Kindern (Kontos & Wilcox-Herzog, 1997; Pianta, La Paro, Payne, Cox & Brad-

ley, 2002). Eine von Feinfühligkeit geprägte Betreuung beeinflusst auch die kognitiven Kompetenzen (Aalsvoort, 2003; Howes & Smith, 1995) sowie die sprachlichen Fähigkeiten von Kindern (McDonald Connor, Son, Hindman & Morrison, 2005) positiv.

Wie bereits in Kapitel 2.1.1 auf theoretischer Basis angenommen, sprechen auch empirische Studien dafür, dass die Selbstbestimmung und Autonomie und somit die aktive Rolle des Kindes von essenzieller Bedeutung für effektive Lernprozesse sind. So zeigten Wahlmöglichkeiten zwischen verschiedenen Aktivitäten (Siraj-Blatchford, Sylva, Mutton, Gilden & Bell, 2002) sowie Verhaltensweisen, die sich am Kind orientierten und weniger direktiv waren (Lerkkanen et al., 2016) positive Zusammenhänge mit der Kompetenzentwicklung der Kinder.

Ein Interaktionsstil, der sich durch Lernunterstützung im Sinne des „scaffolding“ auszeichnet, hat ebenfalls eine positive Wirkung auf die Entwicklung von Kindern (Blewitt et al., 2009; Chien et al., 2010; Pentimonti et al., 2017). Auch Verhalten von Erzieherinnen, welches höhere Denkprozesse aktiviert, korreliert mit besseren Kompetenzen im kognitiven, sprachlichen und sozialen Bereich (z.B. Taylor, Pearson, Peterson & Rodriguez, 2003; Wharton-McDonald, Pressley & Hampston, 1998). Insbesondere Formen des „sustained shared thinking“ scheinen die Kompetenzentwicklung von Kindern positiv zu unterstützen. Aufbauend auf den Daten des EPPE-Projektes betonen Sylva, Melhuish, Sammons, Siraj-Blatchford und Taggart (2004) sowie Siraj-Blatchford et al. (2002) die Schlüsselrolle der Interaktionsform „sustained shared thinking“ für kindliche Lern- und Entwicklungsprozesse. Hildebrandt, Scheidt, Hildebrandt, Hédervári-Heller und Dreier (2016) konnten in einer Studie mit deutschen Kindern positive Effekte von Interaktionen mit Elementen des „sustained shared thinking“ auf das sprachliche Verhalten der Kinder nachweisen.

Ein höherer Umfang von verbalen Interaktionen zwischen Erzieherin und Kind geht mit einer positiven sprachlichen Entwicklung (McCartney, 1984) sowie besseren sozialen Kompetenzen (Clarke-Stewart, 1987) einher. Der Einsatz von offenen Fragen zeigt in Studien zur Erzieherin-Kind-Interaktion einen positiven Zusammenhang mit der verbalen Beteiligung (de Rivera et al., 2005) sowie der Sprachentwicklung der Kinder (Dickinson & Tabors, 2001; Siraj-Blatchford et al., 2002). Der sprachförderliche Effekt von offenen Fragen konnte auch in Studien zum Dialogischen Lesen bestätigt werden (Whitehurst et al., 1994). Darüber hinaus gibt es wissenschaftliche Hinweise darauf, dass eine syntaktisch komplexe und inhaltlich vielfältige Sprache (Vasilyeva, Huttenlocher & Waterfall, 2006) sowie das Wiederholen und Erweitern kindlicher Äußerungen (Girolametto & Weitzman, 2002) sich positiv auf die sprachliche Entwicklung von Kindern auswirken. In verschiedenen Untersuchungen konnte ebenfalls eine bessere sprachliche Anregung mit höheren sprachlichen Fähigkeiten der

Kinder in Verbindung gebracht werden (Connor, Morrison & Slominski, 2006; Girolametto & Weitzman, 2002).

---

Die Ergebnisse wissenschaftlicher Studien belegen, dass die Interaktionen zwischen pädagogischer Fachkraft und Kind einen bedeutenden Einfluss auf die kindliche Entwicklung haben. Hinsichtlich der Stärke des Zusammenhangs merkt Pianta (2017) an, dass es sich insgesamt um eher kleine Zusammenhänge handelt. Etwas größere Effektstärken wurden in Bezug auf Kinder mit einem erhöhten Entwicklungsrisiko gefunden (z.B. Hamre & Pianta, 2005; Ladd & Burgess, 2001).

### Qualität von Fachkraft-Kind-Interaktionen in Kindertageseinrichtungen

Eine große Anzahl an empirischen Studien bestätigt einen signifikanten Zusammenhang des Interaktionsverhaltens von pädagogischen Fachkräften mit kindlichen Entwicklungsmaßen. Qualitätsstudien weisen allerdings insgesamt auf eine mittelmäßige Interaktionsqualität hin (z.B. Hamre et al., 2013; Mashburn et al., 2008; Mayer, Beckh, Berkic & Becker-Stoll, 2013). Die Ausprägung der Qualität unterscheidet sich hinsichtlich der verschiedenen Aspekte von Interaktionen und soll daher im Folgenden etwas differenzierter dargestellt werden.

Die emotionale Unterstützung von pädagogischen Fachkräften befindet sich in internationalen Untersuchungen in der Regel im mittleren bis hohen Qualitätsbereich. Dies gilt sowohl für die Studien, die den „emotional support“ mit der CLASS beobachten (z.B. Mashburn et al., 2008; Pianta, Belsky, Houts & Morrison, 2007) als auch für Untersuchungen, welche andere Instrumente einsetzen, um das emotionale Klima und die Feinfühligkeit von pädagogischen Fachkräften im Kindergarten zu messen (z.B. Helmerhorst, Riksen-Walraven, Fukkink, Tavecchio & Gevers Deynoot-Schaub, 2017). Dieses Ergebnis findet sich auch im deutschsprachigen Raum wieder. In der Studie von Wirts et al. (2017) zu den „Bedingungsfaktoren für gelingende Interaktionen zwischen Erzieherinnen und Kindern“ (BIKE) wies die mit der CLASS erfasste emotionale Unterstützung der Fachkräfte durchschnittlich eine hohe Qualität auf. Vor allem das positive Klima und die Feinfühligkeit zeigten ein hohes Niveau. Übereinstimmend damit stellt König (2009) fest, dass die Atmosphäre in den Kindergärten in der Regel durch ein wertschätzendes Eingehen auf die einzelnen Kinder geprägt ist.

Ein ähnlicher Befund konnte in Bezug auf Verhaltensweisen festgestellt werden, die die Organisation des Alltags und die Steuerung und Regulierung des kindlichen Verhaltens in der Gruppe betreffen. Werte für den CLASS-Bereich „classroom organization“ liegen sowohl in internationalen Studien (z.B. Curby et al., 2009; Rimm-Kaufman et al., 2009) als auch in deutschen Studien (z.B. Suchodoletz, Fäsche, Gunzenhauser & Hamre, 2014; Wirts et al.,



2017) im mittleren bis oberen Bereich der sieben-stufigen Skala. Auch Helmerhorst et al. (2017) konnten beobachten, dass die Strukturierung des Alltags und die verhaltensregulierenden Interaktionen von Fachkräften im Durchschnitt sehr positiv ausfallen.

Ein anderes Bild ergibt sich hinsichtlich lernunterstützender Interaktionen. Im Bereich „instructional support“ der CLASS konnte in zahlreichen amerikanischen Studien nur eine geringe Qualität registriert werden (z.B. Curby et al., 2009; Hamre et al., 2013; Mashburn et al., 2008; Pianta et al., 2007). Mit Mittelwerten zwischen 2 und 3 ist die Qualität der Dimensionen von „instructional support“ in Deutschland ebenfalls im niedrigen bis mittleren Bereich angesiedelt (z.B. Suchodoletz et al., 2014; Wirts et al., 2017). Dies zeigte sich ebenfalls in Studien, welche lernanregende Aspekte im Interaktionsverhalten mit anderen Instrumenten erfassten (z.B. Anders et al., 2012; Helmerhorst et al., 2017; König, 2009).

Die sprachliche Lernunterstützung erreicht im Durchschnitt nur ein niedriges oder knapp mittleres Qualitätsniveau (z.B. Justice, Mashburn, Hamre & Pianta, 2008; Kammermeyer, Roux & Stuck, 2013). Die Zeit mit den Kindern wird oft nur wenig für Gespräche mit den Kindern genutzt (Kontos & Wilcox-Herzog, 1997; König, 2009). Bei den beobachteten Dialogen überwiegen eher Gespräche organisatorischer Art gegenüber inhaltlichen Dialogen (Gasteiger-Klicpera, Knapp & Kucharz, 2010). Auch offene Fragen kommen im Kita-Alltag nur selten vor (Briedigkeit, 2011; Siraj-Blatchford & Manni, 2008). Dabei konnten in Lesesituationen verhältnismäßig mehr offene Fragen beobachtet werden (Massey et al., 2008). Die Gespräche und Aktivitäten im Kindergarten scheinen zudem vor allem durch die pädagogischen Fachkräfte gelenkt zu werden, für Initiativen der Kinder bleibt oft nur wenig Raum (Göncü & Weber, 2000). Strategien, mit denen die Kinder im Sinne des „scaffolding“ in ihrer „Zone der nächsten Entwicklung“ angeregt werden, werden von Fachkräften nur selten eingesetzt (Pentimonti & Justice, 2010; Pentimonti et al., 2017). Darüber hinaus gibt es einige Studien, die zeigen, dass Kinder im Kindergarten nur selten zu höheren Denkprozessen angeregt werden (z.B. Hildebrandt et al., 2016; König, 2009; Siraj-Blatchford et al., 2002).

## 2.2 Möglichkeiten der Qualitätsentwicklung

Die empirischen Befunde zur Interaktionsqualität in Kindertageseinrichtungen deuten auf einen signifikanten Einfluss für die kindliche Entwicklung hin, zeigen aber gleichzeitig auf, dass vor allem das Qualitätsniveau lernunterstützender Interaktionen nicht zufriedenstellend ist. Dieses Kapitel soll sich daher mit der Frage beschäftigen, wie die Qualität der Interaktionen von pädagogischen Fachkräften verbessert werden kann.

### 2.2.1 Ansätze zur Verbesserung der Prozessqualität

Im 12. Kinder- und Jugendbericht des BMFSFJ (2005) werden zwei Wege genannt, über die Verbesserungen in der Prozessqualität von Kindertageseinrichtungen erreicht werden können: indirekt durch eine Verbesserung von relevanten Merkmalen der Struktur- und Orientierungsqualität und direkt durch die Einführung von Qualitätsmanagementsystemen bzw. die Fortbildung der pädagogischen Mitarbeiter in den Einrichtungen.

#### Indirekte Verbesserung über Veränderungen in Struktur- und Orientierungsqualität

Die Ergebnisse von empirischen Studien weisen zum Teil darauf hin, dass bestimmte strukturelle Aspekte einen Einfluss auf die Prozessqualität von Kitas und somit die Interaktionen zwischen Fachkräften und Kindern haben. Vor allem die Ausbildung von pädagogischen Mitarbeitern und der Betreuungsschlüssel werden in diesem Kontext immer wieder genannt (z.B. Howes, Phillips & Whitebook, 1992; NICHD ECCRN, 2002; Siraj-Blatchford et al., 2002). Mit Bezug auf die Qualifikation der Fachkräfte ist in den letzten Jahren ein Prozess der Professionalisierung in der Frühpädagogik in Gang gekommen. Die Anhebung der Ausbildung von Erzieherinnen auf Fachhochschul- bzw. Universitätsniveau stellt dabei einen zentralen Reformimpuls dar. In diesem Zusammenhang kam es zu einer Einführung von konsekutiven und grundständigen Studiengängen der Kindheitspädagogik, welche sich mittlerweile an vielen Fachhochschulen etabliert haben (Altermann & Holmgaard, 2015). Zu den Effekten eines höheren Ausbildungsniveaus auf die Qualität gibt es in Deutschland bisher keine fundierten Erkenntnisse. In den USA gibt es einzelne Studien, die darauf hinweisen, dass ein Bachelor-Abschluss nicht zwangsweise mit einer höheren Qualität einhergeht (z.B. Early et al., 2007). Das BMFSFJ (2005) betont in diesem Zusammenhang, dass nicht nur das formale Niveau, sondern vor allem die Inhalte der Ausbildung in den Blick genommen werden sollten.

Ein durchschnittlicher Betreuungsschlüssel von 1:10 wird als qualitativ akzeptabler Richtwert im Kindergartenbereich angesehen (BMFSFJ, 2005). Da ein Teil der Arbeitszeit von Fachkräften für die Vor- und Nachbereitung der Arbeit verwendet werden muss, empfiehlt die Bertelsmann Stiftung einen Personalschlüssel von 1:7,5 für Kinder ab drei Jahren. Dieser wird in den meisten Bundesländern noch nicht erfüllt, weshalb ein weiterer Aufbau von Stellen nötig ist (Bertelsmann Stiftung, 2015).

In einem gemeinsamen Beschluss der Jugendministerkonferenz/Kultusministerkonferenz (2004) wurden folgende wichtige Bildungsbereiche für die Arbeit mit Kindern in der Kindertagesbetreuung spezifiziert: „Sprache, Schrift, Kommunikation“, „Personale und soziale Entwicklung, Werteerziehung/religiöse Bildung“, „Mathematik, Naturwissenschaft, (Informations-)Technik“, „Musische Bildung/Umgang mit Medien“, „Körper, Bewegung, Gesund-

heit“, „Natur und kulturelle Umwelten“. Des Weiteren wird in dem Bericht auf das Bild vom Kind, die Rolle von Erzieherinnen und Eltern, die räumliche Gestaltung sowie Aspekte der Kooperation und Vernetzung eingegangen. Der Beschluss setzte einen gemeinsamen Rahmen der Orientierungsqualität für die Kindertagesbetreuung in Deutschland. Er wurde mittlerweile von den meisten Bundesländern durch landesspezifische Rahmenpläne ergänzt.

### Direkte Verbesserung der Prozessqualität

Zur Verbesserung der Prozessqualität wurden in den letzten Jahren verschiedene Qualitätsmanagementsysteme eingeführt. Dazu zählt z.B. die Zertifizierung nach der DIN EN ISO 9001:2000. Ziel solcher Zertifizierungen ist es, Bedingungen zu kontrollieren, welche erfüllt sein sollten, damit die Qualität der erbrachten Leistungen erwartet werden kann (Hippel & Grimm, 2010). Die Zertifizierungen sind immer nur für einen begrenzten Zeitraum gültig und beinhalten regelmäßige Überprüfungen (BMFSFJ, 2005).

Bemühungen, die Prozessqualität über Veränderungen der strukturellen Rahmenbedingungen und Orientierungen zu verbessern, sind wichtige Bestandteile der Qualitätsentwicklung, jedoch nicht ausreichend (Early et al., 2007). So stimmen beispielsweise 96,6 % der teilnehmenden Träger einer Befragung zur Akademisierung in der Frühpädagogik der Aussage zu, dass neben der Akademisierung die Qualifizierung des Personals durch Fort- und Weiterbildung von gleicher Bedeutung ist (Altermann & Holmgaard, 2015). Auch eine Verbesserung des Betreuungsschlüssels greift zu kurz, da bei gleichen Rahmenbedingungen noch genügend Spielraum für unterschiedliches Handeln bleibt (BMFSFJ, 2005). Die Fort- und Weiterbildung von pädagogischen Fachkräften stellt daher einen unverzichtbaren Bestandteil bei der Verbesserung von Prozessqualität dar. In den folgenden Abschnitten soll näher auf diese Möglichkeit der Qualitätsentwicklung eingegangen werden.

## **2.2.2 Fort-und Weiterbildung als Säule der Qualitätsentwicklung**

In Deutschland beschäftigt sich die „Weiterbildungsinitiative Frühpädagogische Fachkräfte“ (WiFF) eingehend mit der Professionalisierung von frühpädagogischen Fachkräften im Rahmen von Fort- und Weiterbildungen. Für die Unterscheidung der Begriffe Fortbildung und Weiterbildung gibt es keine klare Definition. Im normalen Sprachgebrauch werden Fortbildungen als etwas thematisch eingegrenztes gesehen, wohingegen Weiterbildungen häufig als Veranstaltungen verstanden werden, die längerfristig angelegt sind und mit einem Abschluss oder Zertifikat beendet werden (Weiterbildungsinitiative Frühpädagogische Fachkräfte, 2011). WiFF verwendet für beide Aspekte den Begriff „Weiterbildung“. In der vorliegenden Arbeit werden beide Begriffe synonym gedacht. Weiterbildung wird im All-

gemeinen als „*Fortsetzung oder Wiederaufnahme organisierten Lernens nach Abschluss einer unterschiedlich ausgedehnten ersten Bildungsphase*“ betrachtet (Deutscher Bildungsrat, 1970, S. 197). Ergänzend dazu umfassen für Tenorth und Tippelt (2007, S. 764) Weiterbildungen „*alle organisierten Maßnahmen, die zum Erhalt, zur Verbesserung oder zum Erwerb neuer Kenntnisse und Qualifikationen dienen*“.

### Weiterbildungslandschaft in der Frühpädagogik

Durch den stärkeren Fokus auf die Qualifizierung von Fachkräften kam es zu einer enormen Expansion des Weiterbildungsmarktes in Deutschland, welche sich in einer Zunahme der Anbieter und der Erweiterung des Angebots äußert (Baumeister & Grieser, 2011). Häufig wird die damit einhergehende zunehmende Heterogenität und die fehlende Transparenz kritisiert (Weiterbildungsinitiative Frühpädagogische Fachkräfte, 2011). Baumeister und Grieser (2011) haben im Rahmen von WiFF eine Analyse der Angebote der großen Weiterbildungsanbieter in Deutschland vorgenommen und diese hinsichtlich Dauer, Methodik und Inhalten beschrieben. 87,7 % der Angebote dauern zwischen einem und drei Tagen, wobei sich fast die Hälfte auf einen Tag begrenzt. Seminare bilden mit 84,6 % den größten Anteil unter den Fortbildungsformaten. In den Veranstaltungen werden zwar vielfältige Methoden eingesetzt, eine Anwendung in der Praxis ist jedoch nur selten Bestandteil. Lediglich bei 4,8 % aller Weiterbildungsangebote findet eine Lernerfolgskontrolle statt. Nach Schmidt (2010) kann eine Qualitätssicherung in der Weiterbildung jedoch nur durch eine systematische Evaluation sichergestellt werden, bei der die Kompetenzentwicklung der Lernenden sowie die Identifizierung von möglichen Schwachstellen im Mittelpunkt steht. Da die Kompetenzen der Lernenden eine wesentliche Rolle im Kontext der Qualifizierung von Fort- und Weiterbildung spielen, soll der Kompetenzbegriff im nächsten Abschnitt näher beleuchtet werden.

### Kompetenzorientierung in der Fort- und Weiterbildung

Die gängigste Definition von Kompetenz stammt von Weinert (2001, S. 27), welcher Kompetenzen als „*die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können*“, betrachtet. Zentral ist dabei die Unterscheidung von Disposition und Performanz. Disposition meint die prinzipielle Fähigkeit, bestimmte Anforderungssituationen zu bewältigen. Performanz umfasst die tatsächlich umgesetzte Handlung, welche von einer situations- und motivationsabhängigen Handlungsbereitschaft bestimmt wird (Nentwig-Gesemann, Fröhlich-Gildhoff & Pietsch, 2011). Es gibt drei verschiedene Arten von Kompetenzmodel-

len. Strukturmodelle unterscheiden verschiedene inhaltliche Teildimensionen von Kompetenz, z.B. Fachkompetenz, Methodenkompetenz, Personalkompetenz und Sozialkompetenz. Stufenmodelle beschreiben unterschiedliche Niveaustufen von Kompetenz in einer bestimmten Domäne. In der Frühpädagogik wird häufig auf Prozessmodelle zurückgegriffen, welche den prozessualen Charakter professionellen Handelns betonen. Eine professionelle Handlung besteht demnach aus einem Kreislauf von Wissen, Analyse, Recherche, Planung, Durchführung sowie Evaluation. Das wohl bekannteste Kompetenzmodell in der Frühpädagogik wurde von Fröhlich-Gildhoff, Nentwig-Gesemann und Pietsch (2011) entwickelt (vgl. Abbildung 2).

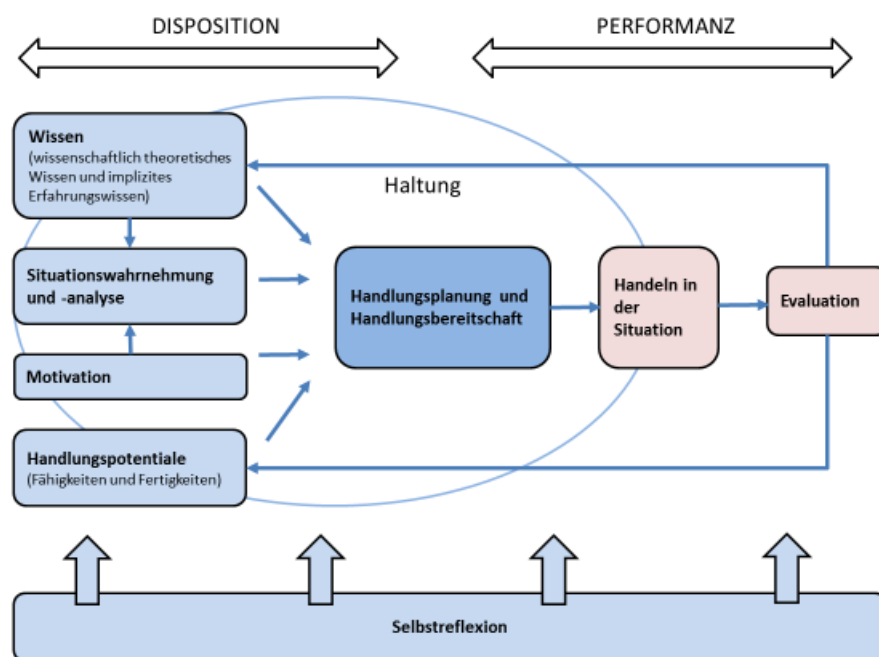


Abbildung 2: Kompetenzmodell nach Fröhlich-Gildhoff et al. (2011)

Das Modell betont die Unterscheidung zwischen Handlungsgrundlagen (Disposition), Handlungsbereitschaft und Handlungsrealisierung (Performanz). Die Grundlage der Handlungsfähigkeit sind das Wissen sowie die Fähigkeiten und Fertigkeiten der pädagogischen Fachkraft. Die Bereitschaft zu einer Handlung resultiert zudem aus einer Situationsanalyse und der Motivation. Die Haltung der jeweiligen Person liegt hinter der Ebene der Disposition und bestimmt wesentlich die Umsetzung der Dispositionen in eine Handlung mit. Das Handeln in der Situation wird schließlich, zumindest implizit, evaluiert und beeinflusst wiederum die Handlungspotenziale. Das entscheidende Ziel bei der Qualifizierung von Fachkräften ist nicht, dass eine Fachkraft im Prinzip kompetent handeln könnte, sondern dass die erworbenen Kompetenzen in der konkreten Alltagspraxis auch situationsangemessen ein- bzw. umgesetzt werden. Lehr-Lern-Arrangements sollten daher vor allem auf die per-

formative Struktur von Kompetenzen abzielen und somit den Transfer des Gelernten in die Praxis unterstützen (Fröhlich-Gildhoff, 2014). Unter Transfer wird in diesem Sinne die *„erfolgreiche Anwendung angeeigneten Wissens bzw. erworbener Fertigkeiten im Rahmen einer neuen, in der Situation der Wissens- bzw. Fertigungsaneignung noch nicht vorgekommenen Anforderung“* (Mähler & Stern, 2006, S. 782) verstanden.

### Wirksamkeit von Fort- und Weiterbildung in der Frühpädagogik

Welche Effekte Fort- und Weiterbildungen auf die Kompetenzen von frühpädagogischen Fachkräften haben können, vor allem auf deren tatsächliche Umsetzung (Performanz), soll an dieser Stelle thematisiert werden. Dazu werden zunächst die Ergebnisse verschiedener Meta-Analysen zu den Effekten von Fort- und Weiterbildung skizziert. Zudem werden einzelne Fort- oder Weiterbildungsformate aus dem frühpädagogischen Feld vorgestellt, die sich als wirksam erwiesen haben. Schließlich sollen Rahmenmodelle dargestellt werden, welche die Bedingungsfaktoren von Weiterbildungseffekten systematisch gebündelt haben.

#### *Ergebnisse von Meta-Analysen*

In den vergangenen 10 Jahren sind vor allem vier bedeutende Meta-Analysen zu den Effekten von Fort- und Weiterbildung im pädagogischen Kontext entstanden. Die Ergebnisse dieser werden im Folgenden der Reihenfolge ihrer Entstehung nach beschrieben.

Fukkink und Lont (2007) geben mit ihrer Meta-Analyse einen Überblick über Studien zu den Trainingseffekten hinsichtlich der Interaktionskompetenzen von Erzieherinnen von Kindertageseinrichtungen, welche zwischen 1980 und 2005 entstanden sind. Sie konnten 15 Studien mit 17 Interventionen und insgesamt 78 Effektstärken einbeziehen. Die Trainings hatten im Durchschnitt eine Effektstärke von  $d = 0.45$ , was nach Cohen (1988) einem mittelgroßen Effekt entspricht. Darin waren Effekte auf das Wissen, die Einstellungen und das Interaktionsverhalten der Fachkräfte enthalten. Die Effekte auf das Verhalten waren mit einer Stärke von  $d = 0.40$  kleiner als die Effekte auf das Wissen ( $d = 0.43$ ) oder die Einstellungen ( $d = 0.65$ ). Die Fortbildungseffekte waren kleiner, wenn nicht mit einem festgelegten Curriculum gearbeitet wurde, die Fortbildung an verschiedenen Orten stattfand und es eine große Anzahl an Teilnehmern gab. Vor allem letzteres deutet darauf hin, dass sogenannte „large-scale“-Programme weniger erfolgreich sind. Die Effekte waren größer, wenn Instrumente verwendet wurden, die auf den Inhalt des Trainings ausgerichtet waren. Bei vier Studien konnten Effektstärken in Bezug auf kindliche Entwicklungsmaße analysiert werden. Es konnten positive Veränderungen in der Kompetenz der Kinder festgestellt werden, welche jedoch nicht signifikant waren. Dies könnte nach Ansicht der Autoren auf die geringe Anzahl an Studien zurückgeführt werden (Fukkink & Lont, 2007).

Durch eine systematische Literaturrecherche konnte Egert (2015) 36 Studien mit insgesamt 42 Interventionen ausfindig machen, die zwischen 1970 und 2011 veröffentlicht wurden und die Effekte von Training auf Qualitätsmaße in der Frühpädagogik untersuchten. Die Meta-Analyse ergab eine durchschnittliche Effektstärke von  $g = 0.68$ , welche ebenfalls als mittelgroß bezeichnet werden kann. Mit dem Effektstärkemaß wurden dabei Verzerrungen aufgrund von kleinen Strichprobengrößen korrigiert (Hedges & Olkin, 1985). Es konnten signifikant größere Trainingseffekte festgestellt werden, wenn renommierte Skalen ( $g = 0.76$ ) gegenüber neu entwickelten Instrumenten ( $g = 0.41$ ) eingesetzt wurden. Unter den bereits etablierten Verfahren erzielten Evaluationen, welche die Interaktionsqualität mit der CLASS gemessen haben, die größten Effekte ( $g = 1.04$ ). Ein linearer Zusammenhang zwischen Trainingsdauer oder -umfang und den Trainingseffekten wurde nicht gefunden. Eine Kategorisierung der Variablen führte zu dem Ergebnis, dass Trainings mit einem Umfang von 45 bis 60 Stunden signifikant höhere Effekte hatten als andere. Um die verschiedenen Formate der Fortbildungsprogramme vergleichen zu können, wurden die eingesetzten Methoden miteinander verglichen, z.B. „workshop“, „coursework“, „coaching“, „online and distance training“ oder „onsite support“. Die Analysen zeigten, dass Trainings, welche Coaching als alleinige Methode einsetzten ( $g = 1.98$ ), fast drei Mal so große Effekte hatten im Vergleich zu anderen Fortbildungsformaten ( $g = 0.67$ ). Mithilfe einer genaueren Analyse der besonders effektiven Trainings konnte Egert (2015) herausfinden, dass diese häufig sehr gut ausgebildete und speziell geschulte Trainer sowie Feedback, aktive Lernelemente sowie praktische Arbeitsaufträge einsetzten. An weiteren 48 Studien wurden die Effekte von Training auf die sprachliche Entwicklung von Kindern untersucht. Mit einer Effektstärke von  $g = 0.38$  handelt es sich hierbei im Durchschnitt um eher kleine Effekte.

In der Meta-Analyse von Werner, Linting, Vermeer und van IJzendoorn (2016) wurden die Effekte von 16 Studien (mit 19 Stichproben) integriert, welche die Wirkung von Interventionen bezüglich der Qualität auf Gruppen-Ebene, auf Fachkraft-Ebene oder auf Kind-Ebene mithilfe eines randomisierten Kontrollgruppen-Designs evaluierten. Der durchschnittliche Effekt über alle Qualitätsmaße hinweg hatte eine Größe von  $g = 0.35$ , wobei die Effekte auf Fachkraft-Ebene ( $g = 0.44$ ) größer waren als die Effekte auf Gruppen-Ebene ( $g = 0.39$ ) oder Kind-Ebene ( $g = 0.26$ ). Eine Moderatoren-Analyse zeigte, dass der Umfang und die Dauer des Trainings keinen Einfluss auf den Effekt haben. Programme, die individuelle Trainingseinheiten umfassten ( $g = 0.41$ ), hatten signifikant höhere Effekte als Trainings ohne individuelle Elemente ( $g = 0.09$ ).

In der aktuellsten Meta-Analyse von Markussen-Brown et al. (2017) wurden anhand von 25 Studien die Effekte von Interventionen mit dem Fokus auf Sprache und Literacy auf die Prozessqualität, die Strukturqualität, das Wissen der Fachkräfte und die sprachliche Kompetenz

der Kinder überprüft. Die Effektstärke wird durch die standardisierte Mittelwert-Differenz (SMD) angegeben, welche durch eine Umwandlung in „Hedges's  $g$ “ korrigiert wurde. Die Effekte der Interventionen auf die Prozessqualität waren von mittlerer Stärke (SMD = 0.59), auf struktureller Ebene konnten sogar große Effekte gefunden werden (SMD = 1.07). Keine signifikante Wirkung hatten die Trainings auf das Wissen der Fachkräfte (SMD = 0.12). Mit Blick auf die kindlichen Entwicklungsmaße ergaben sich kleine Effekte zwischen 0.12 und 0.30. Wie bereits Egert (2015) konnten auch Markussen-Brown et al. (2017) feststellen, dass Formate mit Coaching signifikant höhere Effekte hatten als Interventionen ohne Coaching. Darüber hinaus waren die Effekte umso größer, je länger und intensiver die Interventionen angelegt waren. Dieser Befund widerspricht zum Teil den Ergebnissen von Egert (2015) und Werner et al. (2016).

Betrachtet man die Ergebnisse der verschiedenen Meta-Analysen, ergibt sich folgendes Bild: Fortbildungsmaßnahmen im Bereich der Frühpädagogik haben im Durchschnitt mittelgroße Effekte auf die Qualität der Betreuung. Dabei scheinen Interventionen mit individuellen und situativen Elementen, wie Coaching, besonders wirksam zu sein. Hinsichtlich der Dauer und des Umfangs von Trainings gibt es inkonsistente Befunde. Insgesamt ist zu beachten, dass es zum Teil große Unterschiede in den Effekten zwischen den Studien gibt und immer wieder auch negative Effekte berichtet werden (Egert, 2015). Viele Studien konnten zudem aufgrund von mangelnder Methodik erst gar nicht in die genannten Meta-Analysen einbezogen werden. Im Folgenden werden einzelne Fortbildungsformate vorgestellt, die sich als wirksam erwiesen haben.

### *Wirksame Fortbildungsformate*

Die Forschungsgruppe um Pianta entwickelte „MyTeachingPartner“ (MTP) – ein Ansatz beruflicher Weiterbildung, der darauf abzielt, qualitativ hochwertige Fachkraft-Kind-Interaktionen im Sinne der Qualitätsdimensionen der CLASS zu unterstützen (Pianta, Mashburn, Downer, Hamre & Justice, 2008). Das Programm besteht aus zwei Komponenten: Zum einen erhalten die Fachkräfte Zugang zu Videoaufnahmen, die gelungene Interaktionsbeispiele darstellen. Zum anderen nehmen die Fachkräfte an einem Beratungsprozess teil, der aus einem regelmäßigen, gezielt auf die CLASS-Dimensionen abgestimmten Feedback besteht. Das Feedback wird von geschulten Mitarbeitern nach einem standardisierten Protokoll durchgeführt. Zusätzlich werden ihnen zwei Curricula zur Verfügung gestellt, die sich mit der Unterstützung der sozial-emotionalen und sprachlichen Entwicklung von Kindern auseinandersetzen. Im Rahmen der Beratung machen die teilnehmenden Fachkräfte alle zwei Wochen eine Videoaufnahme von der Umsetzung einer Aktivität aus dem Curriculum und schicken diese an den Trainer. Dieser analysiert das Video und verfasst ein schriftliches Feedback inklusive reflektierender Fragen an die Fachkraft. Auf einem



persönlichen Portal kann sich die Erzieherin dieses Feedback anschauen. Zusätzlich findet ein Gespräch mithilfe eines Video-Chats statt, bei dem sie gemeinsam die einzelnen Interaktionen im Video besprechen und ggf. alternative Handlungsmöglichkeiten diskutieren. In der Untersuchung von Pianta und Mashburn et al. (2008) mit 113 Teilnehmern konnten sich Fachkräfte, die das individuelle Feedback erhielten, signifikant mehr in ihrem Interaktionsverhalten verbessern als Fachkräfte, die nur Zugang zu den „Good practice“-Filmen hatten. Die Unterschiede erzielten eher kleine Effektstärken.

Das sogenannte „Hanen-Program“ mit dem Titel „Learning Language and Loving It“ zielt speziell auf die Unterstützung der sprachlichen Entwicklung von Kindern in Kindertageseinrichtungen ab (Girolametto, Weitzmann & Greenberg, 2003). Das Training wird von einem geschulten und zertifizierten Berater durchgeführt und besteht aus mehreren Gruppensitzungen mit bis zu 20 Teilnehmern und individuellen Sitzungen. In den Gruppensitzungen werden unter anderem Videoaufnahmen mit Praxisbeispielen analysiert sowie Gruppendiskussionen und Rollenspiele durchgeführt. In jeder individuellen Sitzung wird das Interaktionsverhalten der jeweiligen Fachkraft anhand einer fünf-minütigen Videoaufnahme besprochen. Dabei erhält die Fachkraft ein individuelles Feedback durch den Trainer, bei dem es darum geht, wie die im Programm vermittelten sprachförderlichen Strategien (z.B. offene Fragen stellen) umgesetzt wurden. Die positiven Effekte des „Hanen-Program“ auf die sprachliche Anregung von Fachkräften und das daraus resultierende sprachliche Verhalten der Kinder konnten in verschiedenen Studien bestätigt werden (Girolametto et al., 2003; Girolametto, Weitzman, Lefebvre & Greenberg, 2007; Milburn et al., 2014). Die durchschnittlichen Effekte bei Bilderbuchbetrachtungen waren dabei von mittlerer Stärke.

Eine ähnliche Qualifizierung wird in Deutschland im „Heidelberger Interaktionstraining“ (HIT) umgesetzt. Das Programm richtet sich an Fachpersonal im Bereich der Frühpädagogik mit dem Ziel, ihre Kompetenzen in Bezug auf die Unterstützung von sprachauffälligen Kindern in ihrer sprachlichen Entwicklung zu erweitern (Buschmann, Simon, Jooss & Sachse, 2010). Dabei steht das Erlernen von sprachförderlichen Strategien, wie korrektivem Feedback oder sprachanregenden Fragen, in Alltagssituationen im Fokus. Ein Schwerpunkt liegt auf der dialogischen Bilderbuchbetrachtung. Die Fortbildung findet in einer festen Gruppe von maximal 15 Teilnehmern statt und erstreckt sich über fünf Termine, wobei die ersten vier Sitzungen innerhalb von drei bis vier Wochen stattfinden und die fünfte Sitzung nach ca. drei Monaten folgt. Genau wie beim „Hanen-Program“ werden die Inhalte durch Präsentationen, Videobeispiele, Gruppenarbeiten und Rollenspiele vermittelt. Darüber hinaus ist die Video-Supervision ein wichtiger Bestandteil des Trainings. Durch die Supervision in der Gruppe haben die Fachkräfte die Möglichkeit, eine eigene Interaktionssituation mit einem sprachauffälligen Kind zu reflektieren, individuelles Feedback diesbezüglich zu erhalten

und von den Beispielen der anderen zu lernen. In empirischen Untersuchungen wurden positive Auswirkungen des Trainings auf das sprachförderliche Interaktionsverhalten der Erzieherinnen (Simon & Sachse, 2011) sowie die sprachlichen Äußerungen der Kinder (Simon & Sachse, 2013) bei einer Bilderbuchbetrachtung gefunden.

Speziell mit den Effekten eines Trainings zur sprachanregenden Gestaltung von Bilderbuchsituationen haben sich auch Wasik et al. (2006) beschäftigt. Pädagogische Fachkräfte, die an einem neunmonatigen Training mit Workshops und individuellem Coaching teilgenommen haben, stellten im Posttest signifikant mehr offene Fragen. Die beteiligten Kinder zeigten ebenfalls bessere sprachliche Fähigkeiten.

Neben den genannten Beispielen gibt es noch eine Reihe von Fortbildungsformaten, deren positive Effekte in empirischen Studien nachgewiesen werden konnten, auf die an dieser Stelle jedoch nicht näher eingegangen wird (z.B. Beller & Beller, 2009; Kucharz et al., 2015; Powell, Diamond, Burchinal & Koehler, 2010).

### *Bedingungsfaktoren für die Effekte von Fortbildungen*

Verschiedene Autoren haben die Effekte von Fortbildungen und die Bedingungen, unter denen sie auftreten, in Form von Rahmenmodellen systematisiert. Nach Baldwin und Ford (1988) gibt es drei Bedingungsfaktoren für den Transfer von Trainingsinhalten: Charakteristika der zu trainierenden Personen, das Design der Trainingsmaßnahme sowie Rahmenbedingungen des Arbeitsumfelds. Egert (2015) erweitert diesen Ansatz und integriert die wissenschaftlichen Befunde zu Trainingseffekten in ein Prozessmodell. Auf der Trainings-Ebene gibt es nach Egert (2015) drei beeinflussende Variablen: „Who“ (Merkmale von Teilnehmer, Trainer sowie deren Passung), „What“ (inhaltliche Aspekte, wie z.B. Vorhandensein eines Curriculums) und „How“ (z.B. Trainingsformat, Intensität und Dauer). All diese Faktoren beeinflussen, inwiefern sich das Training auf das Wissen und die Fähigkeiten der Fachkräfte auswirkt, was sich schließlich in ihrem Handeln äußert. Dieses steht wiederum für eine bestimmte Qualität der Betreuung, welche Auswirkungen auf die kindliche Entwicklung hat.

---

Die Fort- und Weiterbildung wird als eine wichtige Möglichkeit betrachtet, die Kompetenzentwicklung von Fachkräften zu unterstützen. Oberstes Ziel ist dabei, dass eine Fachkraft nicht nur prinzipiell kompetent handeln könnte, also über die entsprechenden Dispositionen als Handlungsgrundlage verfügt, sondern, dass die erworbenen Kompetenzen auch im Alltag umgesetzt werden (Performanz). Verschiedene Meta-Analysen konnten zeigen, dass das Interaktionsverhalten von Fachkräften durch Weiterbildungen positiv beeinflusst werden kann (z.B. Egert, 2015; Werner et al., 2016). Dabei konnten einige gelungene Beispiele für Trainings mit positiven Wirkungen auf das Interaktionsverhalten von Fachkräften identi-

fiziert werden (z.B. „Heidelberger Interaktionstraining“, „MyTeachingPartner“ oder das „Hannen-Program“). Insgesamt setzen viele Fortbildungen in Deutschland hingegen eher auf traditionelle Formate (z.B. Seminare), wobei deren Wirksamkeit häufig gar nicht evaluiert wird (Baumeister & Grieser, 2011). Die Ergebnisse der Meta-Analysen deuten darauf hin, dass vor allem praxisnahe Formate, wie Coaching und Feedback, besonders wirksam sind (z.B. Egert, 2015). Das folgende Kapitel beleuchtet Feedback als eine Methode der Lernunterstützung.

## 3 Feedback

Nachdem in einem ersten Schritt kurz darauf eingegangen wird, was in der Literatur unter Feedback verstanden wird, sollen unterschiedliche Formen von Feedback beschrieben sowie die Wirkung von Feedback auf das Lernen thematisiert werden. Ziel dabei ist es, die Gelingensbedingungen von effektivem Feedback und dessen Umsetzung im frühpädagogischen Kontext herauszuarbeiten.

### 3.1 Begriffsbestimmung

Der Begriff Feedback ist sowohl in der Wissenschaft als auch im Alltag ein fest verankerter Begriff, der jedoch „vielfältige Bedeutungsnuancen“ aufweist (Fengler, 2010). Jeder hat sowohl schon einmal Feedback bekommen als auch anderen Menschen Feedback gegeben. Abhängig vom jeweiligen Kontext variieren die verschiedenen Begriffsverständnisse in Präzision und Breite und münden daher in unterschiedlichen Inhalten und Formen von Feedbacknachrichten.

#### 3.1.1 Definition von Feedback

Eine Definition von Feedback, welche der ursprünglich kybernetischen und systemtheoretischen Sichtweise entspricht, bietet der Duden (Dudenredaktion, o.J.): Feedback wird hier beschrieben als eine *„zielgerichtete Steuerung eines technischen, biologischen oder sozialen Systems durch Rückmeldung der Ergebnisse, wobei die Eingangsgröße durch Änderung der Ausgangsgröße beeinflusst werden kann“*.

In der Systemforschung und Kybernetik wird Feedback im Kontext von Regelungsprozessen betrachtet. Innerhalb eines Regelkreises wird eine bestimmte Größe (Regelgröße) erfasst, anschließend einer Regeleinrichtung zurückgemeldet und mit einem extern festgelegten Soll-Wert (Führungsgröße) verglichen mit dem Ziel, die Regelgröße an die Führungsgröße anzugleichen. Eine ähnliche Sichtweise findet sich auch in der Lernpsychologie (Kulhavy, 1977), in der Feedback vor allem im Kontext von Instruktion betrachtet wird und Informationen meint, *„die Lernenden nach der Bearbeitung von Lernaufgaben bzw. von Teilen dieser Aufgaben bzgl. ihrer Lösung von einer externen Informationsquelle angeboten werden, mit dem Ziel, eine korrekte Lösung dieser Aufgaben in der aktuellen oder auch in künftigen Lernsituationen zu ermöglichen“* (Narciss, 2006, S. 18).

Wichtig ist dabei, dass Feedback dazu benutzt wird, um die Diskrepanz zwischen einem Ist- und einem Soll-Wert zu minimieren oder zu überwinden. Während diese eindeutige Zuordnung von korrekten Lösungen in der experimentellen Lehr-Lernforschung in der Regel

gut funktioniert, erweist sie sich im kommunikationswissenschaftlichen Kontext oft als nicht möglich. Im therapeutischen Rahmen wird Feedback zum Beispiel häufig dazu eingesetzt, eine neue Perspektive aufzuzeigen und so Einsichten anzuregen, zu denen der Klient sonst nicht gekommen wäre (Claiborn & Goodyear, 2005). Eine allgemeingültigere Definition von Feedback als „*die bewusste Rückmeldung von Informationen an eine Person zu ihrem vorherigen Verhalten*“ (Müller & Ditton, 2014, S. 15) stellt den kleinsten gemeinsamen Nenner der unterschiedlichen fachlichen Perspektiven dar. Wesentliches Merkmal dieser Begriffsbestimmung ist die Annahme, dass es sich nur um Feedback handelt, wenn Informationen intentional zurückgemeldet werden. Des Weiteren beinhaltet sie die Abgrenzung von Feedback als extern dargebotene Information von Verhaltenskonsequenzen, die von einer lernenden Person nicht selbst wahrgenommen werden können.

Wie im Rahmen der Definition von Feedback bereits deutlich wird, werden die Begriffe Feedback und Rückmeldung häufig synonym verwendet. Die oben dargestellten Zugänge zum Feedbackbegriff vereint die Tatsache, dass Feedback stark mit der Leistung bzw. dem Verhalten des Feedback-Rezipienten verknüpft ist. Von Rückmeldung wird jedoch auch oft gesprochen, wenn es in einem distalen Verhältnis zum Verhalten der Adressaten steht (z.B. bei Schulvergleichsarbeiten). Müller und Ditton (2014) schlagen daher vor, die beiden Begriffe auf einem Kontinuum anzuordnen, bei dem der Auflösungsgrad der übermittelten Inhalte entscheidend ist. Feedback umfasst verhaltensnahe Informationen, die sich in der Regel auf eine Person beziehen, wohingegen Rückmeldungen eher verhaltensferne Botschaften meinen, die sich nicht auf eine Person beschränken müssen. Kritisch ist an dieser Stelle anzumerken, dass Müller und Ditton (2014) Feedback mithilfe des Begriffes Rückmeldung definieren, was eine Abgrenzung der beiden Begriffe erschwert. In der vorliegenden Arbeit wird in Anlehnung an diese Differenzierung hauptsächlich der Begriff Feedback verwendet, zum Teil aber auch auf den in diesem Kontext gleich zu deutenden Begriff Rückmeldung zurückgegriffen.

### 3.1.2 Inhalte und Formen von Feedback

Selbst wenn man den Begriff Feedback durch oben genannte Definitionen eingrenzt, so bleibt das Spektrum an möglichen Feedbackinhalten und -formen groß. Ansätze, die zahlreichen Feedbackarten zu klassifizieren gibt es bereits von verschiedenen Autoren (z.B. Kulhavy & Stock, 1989; Mason & Bruning, 2001). Narciss (2006) klassifiziert im Kontext der Instruktionsforschung unterschiedliche Feedbackformen anhand von inhaltlichen Aspekten (vgl. Tabelle 1). Dabei werden eher einfache und stärker elaborierte Feedbackformen unterschieden. Bei einem einfachen Feedback wird eine Person vor allem über die Antwortrichtigkeit (KR), die Anzahl der korrekten Lösungen (KP) oder die richtige Lösung (KCR) in-

formiert. Im Falle von elaborierten Feedbackvarianten erhält der Adressat darüber hinaus auch Hinweise zur Art der Aufgabe (KTC), zu vorhandenen Fachbegriffen (KC), zu seinen Fehlern (KM), zur Lösung (KH) oder zu seinem eigenen Lernprozess (KMC).

Tabelle 1: Klassifikation unterschiedlicher Feedbackformen nach Narciss (2006, S. 23)

	Bezeichnung	Beispiele für Feedbackinhalte
Einfache Formen	Knowledge of result / response (KR)	Falsch / richtig Nicht richtig / stimmt
	Knowledge of performance (KP)	15 von 20 Aufgaben richtig 85 % der Aufgaben korrekt gelöst
	Knowledge of correct result (KCR)	Angabe der korrekten Lösung Markierung der korrekten Antwort
Elaborierte Formen	Knowledge on task constraints (KTC)	Hinweise auf Art der Aufgabe Hinweise auf Bearbeitungsregeln Hinweise auf Teilaufgaben Hinweise auf Aufgabenanforderungen
	Knowledge about concepts (KC)	Hinweise auf Fachbegriffe Beispiele für Begriffe Hinweise auf Begriffskontext Erklärungen zu Begriffen
	Knowledge about mistakes (KM)	Anzahl der Fehler Ort der Fehler / des Fehlers Art der Fehler / des Fehlers Ursache/n des/r Fehler(s)
	Knowledge on how to proceed (KH; „know how“)	Fehlerspezifische Korrekturhinweise Aufgabenspezifische Lösungshinweise Hinweise auf Lösungsstrategien Leitfragen Lösungsbeispiele
	Knowledge on meta-cognition (KMC)	Hinweise auf metakognitive Strategien Metakognitive Leitfragen

Die verwendete Terminologie findet vor allem in computer-basierten oder schriftlichen experimentellen Lernumgebungen Anwendung, kann aber auch auf interpersonale Feedbacksituationen übertragen werden. In diesem Fall erfolgen die Übergänge zwischen den einzelnen Formen jedoch eher fließend und sind daher weniger trennscharf (Müller & Dittton, 2014).

Eine spezielle Form von Feedback stellen Multi-Source-Feedbackverfahren dar, bei denen eine Person Feedback von mehreren Quellen erhält. Diese werden zunehmend in der be-

trieblichen Personalentwicklung eingesetzt, um einen genaueren und zuverlässigeren Blick auf das Verhalten der Mitarbeiter zu erhalten (Brutus et al., 2006). Eine populäre Form dieser Verfahren ist das sogenannte 360°-Feedback. In Anlehnung an die eindeutige numerische Bezeichnung schlagen Müller und Ditton (2014) vor, dass Vorgesetzte, Kollegen und Mitarbeiter jeweils 90° in das Feedback einbringen. Dieses 270°-Feedback wird durch eine Selbsteinschätzung des Mitarbeiters ergänzt. Übertragen auf das Schulsystem enthält ein 360°-Feedback für eine Lehrkraft eine Selbsteinschätzung sowie Feedback von Vorgesetzten, Kollegen und Schülern. Kontrovers diskutiert wird dabei die oben aufgeführte Annahme der höheren Genauigkeit von multiperspektivischen Verfahren. In einem Überblick verschiedener Meta-Analysen konstatieren Murphy, Cleveland und Mohler (2001), dass die verschiedenen Beobachter beim 360°-Feedback mit einer Interraterkorrelation von  $r = 0.30$  bis  $0.50$  nicht in der Lage sind, übereinstimmend zu beurteilen. Aus wissenschaftlicher Sicht ist dies eine Messungenauigkeit und wird als problematisch angesehen. Innerhalb der Praxis wird es zum Teil hingegen als bereichernd im Sinne von verschiedenen Perspektiven auf eine Person betrachtet. Hensel, Meijers, van der Leeden und Kessels (2010) stellten fest, dass die Reliabilität der Beurteilungen mit steigender Zahl der Beurteiler höher wurde. Für eine ausreichend hohe Reliabilität von  $r = 0.70$  waren 10 Beurteiler für die Einschätzung von Kompetenzen notwendig, eine reliable Beurteilung der Motivation wurde bei 6 Personen erreicht. Hensel et al. (2010) führen daher die Überlegung an, 360°-Feedback entweder nur für Personalentwicklungsmaßnahmen und nicht im Kontext von Selektion einzusetzen oder die entsprechend nötige Anzahl an Personen einzubeziehen.

## 3.2 Wirkungen und Bedingungen von Feedback

Im Mittelpunkt dieses Kapitels stehen die verschiedenen Wirkungsweisen von Feedback. Die Effekte von Feedback und Gelingensbedingungen werden auf Basis empirischer Erkenntnisse und theoretischer Modelle beleuchtet.

### 3.2.1 Funktionsweisen von Feedback

Wie mit Blick auf die Multi-Source-Feedbackverfahren bereits deutlich wurde, kann Feedback nicht nur auf individueller Ebene greifen, sondern auch eine organisationale Funktion in Form von Personalauswahl oder -entwicklung erfüllen. Da die vorliegende Arbeit nicht im Bereich des „human resource managements“ angesiedelt ist, wird im Weiteren die organisationale Funktion von Feedback vernachlässigt und der Fokus auf die Funktionen von Feedback auf Individualebene gelegt.

In der Literatur gibt es eine breite Darstellung von potenziellen Feedback-Funktionen. Narciss (2006) hat die Differenzierungen verschiedener Autoren zusammenfassend analysiert (vgl. Tabelle 2) und dabei häufig bestätigende/verstärkende, informierende, lenkende/hinweisende sowie regulierende/korrigierende Funktionen identifiziert.

Tabelle 2: Übersicht über postulierte Feedback-Funktionen nach Narciss (2006, S. 77)

Cusella (1987)	Sales (1993)	Wager & Mory (1993)	Butler & Winne (1995)
Verstärken	Stimulieren	Bestätigen	Bestätigen
Informieren	Informieren	Informieren	Informieren
Hinweisen	Lenken	Hinweisen	Hinweisen
Motivieren	Motivieren	Motivieren	
		Korrigieren	Korrigieren
Regulieren	Regulieren		
Instruieren	Instruieren	Instruieren	
	Bewerten	Bewerten	
	Beraten		Anregen
			Ergänzen
			Diskriminieren
			Restrukturieren

Seit den Anfängen der Feedbackforschung werden vor allem drei verschiedene Wirkungsweisen von Feedback postuliert, die dem heutigen Verständnis von Feedback noch sehr nah kommen und sich auch in der oben dargestellten Übersicht wiederfinden. Kulhavy und Wager (1993, S. 5) bezeichnen sie als „feedback triad“: Demnach wird Feedback einerseits als Ansporn/Antrieb betrachtet, welcher dazu dient, die Antwortrate bzw. Antwortrichtigkeit zu erhöhen bzw. ein bestimmtes Ziel zu erreichen. Andererseits wird betont, dass Feedback über die Verknüpfung mit einem Reiz als eine Art Verstärker von korrekten Antworten wirkt. Des Weiteren liefert Feedback Informationen, die der Lernende dazu nutzen kann, um eine vorherige Antwort zu validieren oder anzupassen. Im Wandel der Forschung fanden diese Wirkungswege unterschiedliche Beachtung, was im Folgenden anhand eines kurzen historischen Abrisses skizziert werden soll.

### Feedback als Verstärker

Der Ausgangspunkt der (behavioristischen) Feedbackforschung waren Untersuchungen zum operanten Lernen. Ein wesentliches Prinzip der operanten Lerntheorien stellt dabei das „Law of Effect“ dar (Thorndike, 1920). Dies besagt, dass ein Verhalten dann wiederholt wird, wenn es eine befriedigende Konsequenz nach sich zieht („satisfying state of affairs“). Ebenfalls in den 20er Jahren und inspiriert vom „Law of Effect“ entwickelte Pressey die ersten „Teaching Machines“ (Pressey, 1926; zitiert nach Skinner, 1958). Diese waren ursprüng-



lich für die Intelligenztestung gedacht, er ging jedoch auch von instruktionellen Effekten im Sinne des operanten Lernens aus. Der Lernende wählte mithilfe der Maschine eine Multiple-Choice-Antwort aus. Wenn die Antwort richtig war, ging das Gerät automatisch zum nächsten Item weiter, eine falsche Antwort wurde gezählt und der Lernende musste solange weitermachen bis die richtige Antwort ausgewählt wurde. Pressey war somit der Erste, der die Bedeutung von sofortigem Feedback beim Lernen betonte (Skinner, 1958). Skinner (1968) führte die Arbeit von Pressey mit seinem Instruktionsansatz zum programmierten Unterricht („programmed instruction“) fort. Er ging davon aus, dass der Lernprozess in kleine Schritte unterteilt werden muss, wenn komplexes Verhalten erlernt werden soll. Skinner (1958) entwickelte seine eigene „Teaching Machine“ und nahm dabei vor allem zwei Anpassungen vor. Da die Lernenden die richtige Antwort nicht erkennen, sondern wiedergeben sollen, verwendete er für seine „Teaching Machines“ offene Antworten anstatt des Multiple-Choice-Formats. Bei seinen Geräten verfasste der Lernende eine Antwort, indem er gedruckte Buchstaben oder Zahlen verschob. Die Maschine verglich die Antwort mit einer vorher kodierte richtigen Lösung. Wenn die Antworten übereinstimmten, erschien die nächste Aufgabe - bei einer falschen Antwort wurde diese gelöscht und eine neue Antwort musste verfasst werden. Seinem instruktionellen Ansatz entsprechend, teilte er zudem das Lernmaterial in kleine Untereinheiten auf. Die einzelnen Lernschritte bauen dabei aufeinander auf und gewinnen an Komplexität. Der Lernende konnte nicht zum nächsten Lernschritt weitergehen, bis der vorherige richtig gelöst wurde. Durch dieses Vorgehen verstärkte das Gerät richtige Antworten und versuchte die Auftretenswahrscheinlichkeit von falschen Antworten durch Ignorieren zu verringern, was das Grundprinzip des operanten Lernens darstellt. Skinners Arbeiten stießen in den 50er und 60er Jahren des 20. Jahrhunderts auf so viel Anklang, dass die Perspektive von Feedback als Verstärker zunächst im Mittelpunkt der Forschung stand (Narciss, 2006).

### Feedback als Informationsquelle

Anderson, Kulhavy und Andre (1971) fanden schließlich in ihrer Studie Hinweise darauf, dass Feedback nicht nur korrekte Antworten verstärken, sondern auch falsche Antworten korrigieren kann. Die Gruppe, welche Feedback nach falschen Antworten bekam, zeigte eine höhere Leistung als die Gruppe mit Feedback bei richtigen Antworten. Trotz nicht signifikanter Ergebnisse fand daraufhin ein perspektivischer Wechsel in der Feedback-Forschung statt. Die kognitivistische Perspektive betrachtet vor allem die Prozesse, die zwischen Reiz und Reaktion liegen. Im Zuge solcher Informationsverarbeitungsprozesse werden Informationen vom Lernenden aktiv verarbeitet und im Gedächtnis abgespeichert. Fehler werden dabei als notwendiger Bestandteil des Lernprozesses betrachtet und nicht mehr ignoriert, wie es in der behavioristischen Forschung der Fall war. Im Kontext von kog-

nitiven Ansätzen wird Feedback daher nicht als Verstärker von korrekten Lösungen, sondern als Informationsquelle für die Korrektur von Fehlern gesehen (Krause, 2007). Feedback dient demnach dazu, Verhalten an einen bestimmten Zielzustand anzupassen. Im „TOTE“-Modell (Test-Operate-Test-Exit) von Miller, Galanter und Pribram (1960) findet zunächst eine Art Eingangsdiagnose statt, bei der eine bestimmte Kennzahl mit einem festgelegten Standard abgeglichen wird („Test“). Wird dabei eine Diskrepanz festgestellt, findet eine Regulierung des Verhaltens statt („Operate“), deren Ergebnis wiederum überprüft wird („Test“). Dieser Prozess findet solange statt bis der aktuelle Wert dem Sollwert entspricht, woraufhin die Handlung beendet wird („Exit“). Feedback kann in diesem Sinne definiert werden als *„information about the gap between the actual level and the reference level of a system parameter which is used to alter the gap in some way“* (Ramaprasad, 1983, S. 4). Die Kontrolltheorie von Carver und Scheier (1981) beschreibt mithilfe von drei Ziel-Ebenen (selbstbezogene Ziele, allgemeine Handlungsziele und konkrete Handlungsabläufe) etwas differenzierter ebenfalls ein Feedbacksystem mit Rückkopplungsschleifen. Auf diesen drei hierarchisch angeordneten Ebenen findet ein ständiger Vergleich zwischen Ist und Soll statt. Die Ziele auf den unteren Ebenen werden durch Ziele höherer Ebenen bestimmt. Kopp und Mandl (2014) führen „Fürsorge“ als ein mögliches Beispiel für ein selbstbezogenes Ziel („Be“-goal) an. Dieses oberste Ziel führt nun zum Handlungsziel „Abendessen zubereiten“, welches wiederum in der konkreten Handlung „Brot schneiden“ mündet. Wenn nun das negative Feedback kommt, dass das Brot nicht ausreicht, werden weitere Handlungen ausgelöst (z.B. Brot kaufen).

### Feedback als Motivationsquelle

Die motivierende Funktion von Feedback wird in der Literatur häufig mit der verstärkenden Funktion gleichgesetzt (z.B. Sassenrath, 1975). Erkenntnisse aus Banduras „Self-Efficacy“-Theorie (Bandura, 1977), der Attributionstheorie (Weiner, 1994) sowie der Selbstbestimmungstheorie von Deci und Ryan (1993) weisen darauf hin, dass ein positives Kompetenzerleben als Anreiz für die Erreichung eines Ziels dienen kann. Da jedes Feedback Informationen über Erfolg oder Misserfolg liefert, trägt es zum Kompetenzerleben von Personen bei und kann gegebenenfalls auf diesem Weg die Lernmotivation erhöhen (Narciss, 2006). Eine Gleichsetzung der motivierenden Funktion mit der reinen Verstärkung von richtigen Antworten geht daher nicht weit genug. Die verschiedenen motivationalen Ansätze sollen im Folgenden kurz skizziert werden.

Die im Selbstkonzept verankerten Erwartungen über Handlungskompetenzen und -konsequenzen bezeichnet Bandura (1977) als Selbstwirksamkeitserwartungen. Eine hohe Selbstwirksamkeitserwartung äußert sich in der Überzeugung, über die notwendigen Ressourcen zur Bewältigung von Herausforderungen zu verfügen. Nach seiner Theorie

werden Menschen zu bestimmtem Verhalten motiviert, wenn sie sicher sind, dass sie die entsprechende Handlung ausführen können und diese Handlung auch zum angestrebten Ergebnis führt. Wichtig ist dabei nicht die tatsächliche Kompetenz, sondern die wahrgenommene Selbstwirksamkeit (Bandura, 1993). Selbstwirksamkeitserwartungen sind dynamisch und verändern sich aufgrund von neuen Informationen und Erfahrungen. Da ein Feedback Informationen über das eigene Wissen oder die Fähigkeiten liefert, liegt die Vermutung nahe, dass es die Selbstwirksamkeitserwartungen dieser Person beeinflussen kann.

Wichtige Voraussetzung ist dabei, dass der Erfolg wahrgenommen und richtig attribuiert wird. Handlungsergebnisse in leistungsbezogenen Kontexten können auf eine Vielzahl von verschiedenen Gründen zurückgeführt werden. Weiner (1994) klassifiziert in seiner Attributionstheorie diese Ursachenzuschreibungen auf drei Ebenen. Die Ursachen von Ereignissen können entweder innerhalb (internal) oder außerhalb (external) der eigenen Person lokalisiert (Lokationsdimension), als über die Zeit stabil oder instabil (Dimension der Stabilität) wahrgenommen und als kontrollierbar/veränderbar oder unkontrollierbar/unveränderbar (Dimension der Kontrollierbarkeit) empfunden werden. Weiner (1994) geht davon aus, dass Kausalattributionen die Zielerwartungen und Emotionen und darüber auch das Leistungsverhalten einer Person beeinflussen können. Wenn die Ursache für eine Leistung als internal, instabil und veränderbar bewertet wird (z.B. Anstrengung), so führt dies zu einem intensiverem Leistungsverhalten und einer höheren Ausdauer.

Deci und Ryan (1993) betrachten in ihrer Selbstbestimmungstheorie ebenfalls den Zusammenhang zwischen Kompetenzerleben und Motivation. Sie unterscheiden zwischen intrinsischer und extrinsischer Motivation. Intrinsisch motivierte Verhaltensweisen entstehen vor allem aus persönlichem Interesse und erfordern keine äußeren Anreize. Im Gegensatz dazu drückt sich extrinsische Motivation in Handlungen aus, *„die mit instrumenteller Absicht durchgeführt werden, um eine von der Handlung separierbare Konsequenz zu erlangen“* (Deci & Ryan, 1993, S. 225). Die Selbstbestimmungstheorie betont drei angeborene psychologische Grundbedürfnisse, die für die intrinsische und extrinsische Motivation von Bedeutung sind: Bedürfnis nach Kompetenz oder Wirksamkeit, nach Autonomie oder Selbstbestimmung sowie nach sozialer Eingebundenheit. Intrinsische Motivation ist vor allem mit den Bedürfnissen nach Kompetenz und Selbstbestimmung verknüpft. Nur wenn Lernende sich als kompetent und selbstbestimmt erleben, entwickeln sie eine intrinsische Motivation als Antriebskraft für erfolgreiche Lernprozesse.

### 3.2.2 Empirische Befunde zu den Funktionsweisen von Feedback

Die in Kapitel 3.2.1 postulierten Funktionen von Feedback (verstärkend, informativ und motivierend) wurden in zahlreichen Untersuchungen empirisch überprüft und analysiert. Die Ergebnisse sollen an dieser Stelle exemplarisch anhand von einigen Beispielen beschrieben werden.

#### Studien zur verstärkenden Funktion von Feedback

In wissenschaftlichen Studien zu programmierten Unterrichtssequenzen konnten häufig keine positiven Effekte von ergebnisorientiertem Feedback gefunden werden. Lublin (1965) führte eine Studie mit 219 Studenten des Einführungskurses in Psychologie durch, bei der den Probanden das Lernmaterial im Sinne des programmierten Unterrichts in 26 Teilschritten dargeboten wurde. Die Teilnehmer wurden zufällig einer von vier Bedingungen zugewiesen: a) „continuous reinforcement“ (korrekte Lösung nach jeder Einheit), b) „variable-ratio 50 % reinforcement“ (korrekte Lösung nach 50 % zufällig ausgewählten Einheiten), c) „fixed-ratio 50 % reinforcement“ (korrekte Lösung nach jeder zweiten Einheit) und d) „no reinforcement“ (alle Antworten ausgelassen). Zwischen den Gruppen zeigten sich keine signifikanten Leistungsunterschiede. Ähnliche Ergebnisse finden sich auch in anderen Studien zu verstärkendem Feedback in programmierten Unterrichtseinheiten (z.B. Krumboltz & Weisman, 1962). Im Gegensatz dazu wurden bei einigen Untersuchungen in herkömmlichen Unterrichtssituationen positive Effekte von Feedback gefunden. Mit 123 Studenten der pädagogischen Psychologie untersuchte beispielsweise Cameron (1966) die Wirkung von im Unterricht integriertem Feedback. Er stellte fest, dass die Gruppe, welche nach Antworten im Unterricht ein Feedback erhielt, eine bessere Leistung im Test zeigte als die Gruppe ohne Feedback und schlussfolgerte daraus, dass die frühe Bestätigung einer Antwort die Wahrscheinlichkeit erhöht, dass diese Information erlernt wird.

#### Studien zur informierenden/korrigierenden Funktion von Feedback

Wie bereits erwähnt, fand nach der Studie von Anderson et al. (1971) eine Art kognitive Wende in der Feedbackforschung statt. Es wurde daraufhin in zahlreichen Untersuchungen überprüft, ob Feedback eher eine verstärkende oder eine informative Funktion erfüllt (z.B. Bardwell, 1981; Kulhavy & Anderson, 1972; Sassenrath, 1975). In den drei genannten Untersuchungen wurde die Leistung nach einem „delayed-feedback“ mit der Leistung nach einem „immediate-feedback“ verglichen. Bei Kulhavy und Anderson (1972) sowie Sassenrath (1975) wurde zudem das Antwortverhalten im Posttest differenziert beobachtet, wenn die ursprüngliche Antwort richtig oder falsch war. Ein wesentlicher Bestandteil der operanten Lerntheorien ist das Kontiguitätsprinzip, nach dem es für das Lernen zentral ist, korrekte

Antworten mithilfe von Feedback zeitlich möglichst unmittelbar zu verstärken. Ausgehend von einer verstärkenden Funktion von Feedback müsste demnach bei einer richtigen Antwort im ersten Test die Wahrscheinlichkeit einer richtigen Antwort im Posttest in der Gruppe mit sofortigem Feedback höher sein im Vergleich zur Gruppe mit verzögertem Feedback. Dies konnte in keiner der beiden Untersuchungen bestätigt werden. In der Studie von Bardwell (1981) waren Schüler mit „delayed-feedback“ sowohl beim Erlernen als auch beim Behalten von Vokabeln der Gruppe mit „immediate-feedback“ überlegen. Das Phänomen der besseren langfristigen Leistungen bei verzögertem Feedback gegenüber sofortigem Feedback wird in der Literatur als „delay-retention effect“ bezeichnet. Die Befunde zum „delay-retention effect“ wurden als Widerspruch zur „Verstärkungs-Hypothese“ von Feedback interpretiert (Kulhavy, 1977). Kulhavy und Anderson (1972) führten als mögliche Erklärung die sogenannte „interference-perseveration hypothesis“ an. Nach dieser Theorie vergessen Lernende bei einem verzögertem Feedback falsche Antworten in der Phase zwischen der ersten Antwort und dem Feedback („delay interval“) und es kommt daraufhin zu weniger Störungen („interferences“), wenn die korrekte Antwort durch das Feedback erlernt wird. Wenn das Feedback hingegen unmittelbar erfolgt, wird die falsche Antwort nicht so leicht vergessen („perseverate“) und es kommt leichter zu einer Art Konkurrenz zwischen den beiden möglichen Antworten. Im Falle einer falschen Antwort beim ersten Test müsste demnach ein verzögertes Feedback eher zu einer Korrektur des Fehlers im zweiten Test führen als ein sofortiges Feedback. Sowohl in der Studie von Sassenrath (1975) als auch in der von Kulhavy und Anderson (1972) konnte dies belegt werden. In beiden Fällen war die Wahrscheinlichkeit für die Wiederholung eines ursprünglichen Fehlers im Posttest höher in der Gruppe mit unmittelbarem Feedback als in der Gruppe mit verzögertem Feedback.

### Studien zur motivierenden Funktion von Feedback

Bereits zu Beginn der Feedbackforschung wurde neben den anderen Funktionen auch die motivationale Funktion von Feedback untersucht. In einigen Untersuchungen, in denen Schüler nur über die Anzahl korrekter Lösungen, aber nicht über die richtigen Lösungen noch darüber, welche Aufgaben sie richtig oder falsch beantwortet haben, informiert wurden, wurde trotzdem ein signifikanter Lernzuwachs festgestellt. Die Ergebnisse können somit nicht auf eine verstärkende oder korrigierende Wirkung des Feedbacks zurückgeführt werden. Es kann daher davon ausgegangen werden, dass das Feedback in Form eines „knowledge of performance“ (KP) motivierend wirken und somit das Lernen beeinflussen kann. Auch andere Studien liefern Hinweise auf diesen Wirkmechanismus (z.B. Elwell & Grindley, 1938; Hundal, 1969).

Zur Überprüfung theoretischer Ansätze zur Erklärung der motivierenden Funktion haben ebenfalls einige Untersuchungen stattgefunden. Signifikante Effekte von Feedback auf die

Selbstwirksamkeit von Grundschulern konnten Schunk (1983) sowie Schunk und Swartz (1993) zeigen. In einer aktuellen Studie von Dahling und Ruppel (2016) mit 181 Psychologiestudenten konnten diese Ergebnisse repliziert werden. Narciss (2004) konnte zudem in einem Experiment mit 186 Studenten den signifikanten Einfluss von Feedback und Selbstwirksamkeit auf die Leistung und die Motivation bei einer computer-basierten Lernumgebung zu geometrischen Formen belegen. Angesichts der genannten Forschungsergebnisse kann davon ausgegangen werden, dass Feedback die Selbstwirksamkeit von Lernenden bedeutend beeinflussen kann, was sich wiederum auf die Leistung auswirkt.

In einigen Studien wurde die oben beschriebene Annahme, dass Attributionen von Leistungen eine wesentliche Bedeutung für die Selbstwirksamkeitserwartungen einer Person und somit ihre Leistung haben, untersucht. Dafür wurde häufig die Attribuierung der Leistung im Feedback beeinflusst. Bei Schunk (1983) hatten die Gruppen, die Feedback mit attributionalen Hinweisen zu ihrer Fähigkeit und/oder ihrer Anstrengung erhielten, einen signifikant höheren Zuwachs in den Testwerten und Selbstwirksamkeitswertungen als die Gruppen ohne attributionales Feedback. Zu dieser Erkenntnis sind auch Relich, Debus und Walker (1986) in ihrer Untersuchung mit 84 Schülern der sechsten Klasse in Bezug auf Divisionsaufgaben gekommen.

Sansone (1986) erforschte in zwei Studien mit Psychologiestudenten den Zusammenhang zwischen Feedback und intrinsischer Motivation. Sie stellte fest, dass positives Feedback die wahrgenommene Kompetenz steigern kann und ein hohes Kompetenzerleben schließlich die intrinsische Motivation erhöht. Zu einem ähnlichen Schluss kamen auch andere Studien (z.B. Butler, 1987). In der Meta-Analyse von Deci, Koestner und Ryan (1999) führte Feedback, welches in einer kontrollierenden Art und Weise übermittelt wurde zu einer signifikant niedrigeren intrinsischen Motivation im Vergleich zu informativem Feedback, welches einen motivationsfördernden Effekt hatte. Dies bestätigt die Theorie von Deci und Ryan (1993), dass Lernende dann intrinsisch motiviert sind, wenn sie sich als kompetent und selbstbestimmt erleben.

---

Da in Feedbackuntersuchungen häufig nur eine theoretische Perspektive berücksichtigt wird, werden die verschiedenen Funktionsweisen oft als sich ausschließend betrachtet. Narciss (2006) betont in diesem Zusammenhang, dass es sich eher um ergänzende Funktionen von Feedback handelt. Sie nimmt an, dass eine positive Wirkung von Feedback auf der Verstärkung von richtigen Antworten, der Korrektur falscher Lösungen sowie motivationalen Prozessen basiert. In den vergangenen Abschnitten konnten wissenschaftliche Belege für jede der drei Funktionen aufgezeigt werden. Was jedoch vor allem zu Beginn der Feedbackforschung häufig ignoriert wurde, ist die Tatsache, dass es auch viele andere Stu-

dien gibt, die nicht nur einseitig positive Effekte von Feedback festgestellt haben. In seinem Überblick zu den Effekten von Experimenten mit „knowledge of performance“ (KP) oder „knowledge of result“ (KR) stellt Ammons (1956) beispielsweise die generelle Hypothese auf, dass diese Formen von Feedback das Lernen und die Motivation in den unterschiedlichsten Kontexten positiv beeinflussen, ohne dabei auf widersprüchliche Ergebnisse hinzuweisen (wie z.B. bei Crafts & Gilbert, 1935 oder Deputy, 1929). Darüber hinaus weisen speziell Studien aus den frühen Jahren des letzten Jahrhunderts zum Teil gravierende methodische Mängel in Form von sehr kleinen Stichproben, ungenauen Feedback-Operationalisierungen oder fehlenden Kontrollgruppen auf (Kluger & DeNisi, 1996).

### 3.2.3 Zusammenfassende Ergebnisse von Meta-Analysen

Die Vielzahl an, zum Teil widersprüchlichen, Befunden der empirischen Untersuchungen zu den Effekten von Feedback wurde bereits von einigen Autoren in Form von Meta-Analysen zusammenfassend dargestellt. Zu denen bekanntesten Meta-Analysen gehören die von Bangert-Drowns, Kulik, Kulik und Morgan (1991), Kluger und DeNisi (1996) sowie die aktuellste von Hattie und Timperley (2007).

Bangert-Drowns et al. (1991) haben für ihre Meta-Analyse in der ERIC Datenbank unter den Stichworten „feedback“, „reinforcement“, „knowledge of results“ sowie „instruction“ Studien zu Feedback-Effekten in testähnlichen Lernsituationen ausgewählt. Voraussetzung für den Einbezug in die Analysen war, dass die Studien a) ausreichend empirische Information für die Berechnung von Effektstärken bereitstellten, b) eine Experimentalgruppe mit Feedback mit einer Kontrollgruppe ohne Feedback verglichen, c) frei zugänglich waren und d) keine ernsthaften methodischen Schwachstellen aufwiesen. Insgesamt konnten sie auf diesem Weg 40 Studien von 1960 bis 1990 einbeziehen, welche von zwei Ratern in Bezug auf 16 Merkmale aus den folgenden vier Kategorien ausgewertet wurden: Art des Feedbacks, Forschungsmethode, Untersuchungsaufbau, Veröffentlichungsmerkmale. Der Großteil der Studien wurde mit College-Schülern durchgeführt, wobei die Stichprobengröße in der Regel kleiner als 100 war. Aus den 40 Untersuchungen konnten 58 Effektstärken ermittelt werden. Mit einer mittleren Effektstärke von 0.26 hatten die Feedback-Interventionen im Durchschnitt positive, aber kleine Effekte auf die Leistung der Schüler. Dabei war die Spannweite der Effektstärken recht groß (-0.83 bis 1.42). Von den 58 Effekten waren 18 negativ, wovon wiederum vier signifikant waren.

Mithilfe einer Recherche in verschiedenen Datenbanken identifizierten Kluger und DeNisi (1996) wissenschaftliche Beiträge anhand der Schlüsselwörter „feedback“ oder „KR“ sowie „performance“. Die ca. 3000 ausgewählten Arbeiten wurden anhand der folgenden Kriterien selektiert: Untersuchungsdesign mit mindestens einer Experimentalgruppe mit Feed-

back-Intervention ohne anderes Treatment und mindestens einer Kontrollgruppe, Messung von Leistung, Stichprobengröße von  $N \geq 10$ , die notwendigen Informationen zur Berechnung der Effektstärke  $d$  sind bekannt. Fünf Prozent der Beiträge haben diese Voraussetzungen erfüllt, was einer Anzahl von  $N = 131$  Studien entspricht. Im Durchschnitt haben 39 Personen an der Untersuchung teilgenommen. Im Mittel ergab sich eine an der Stichprobengröße gewichtete Effektstärke von  $d = 0.41$ . Von den 607 errechneten Effektstärken waren über 38 % negativ. Die von den Autoren berichteten Ergebnisse entsprechen daher im Großen und Ganzen den Erkenntnissen von Bangert-Drowns et al. (1991).

Aus einer Synthese von über 500 Meta-Analysen zu den Einflussfaktoren der Leistung von Schülern (Hattie, 1999) haben Hattie und Timperley (2007) 12 Meta-Analysen identifiziert, welche Effekte von Feedback gemessen haben. Diese 12 Meta-Analysen haben insgesamt 196 Studien einbezogen und fast 7000 Effektstärken berechnet. Die Autoren berichten eine mittlere Effektstärke von  $d = 0.79$ , wobei auch hier große Schwankungen vorliegen.

Zusammenfassend kann festgehalten werden, dass Feedback unbestritten das Lernen und die Leistung beeinflussen kann. Allerdings weisen die Studien eine sehr große Varianz in ihren Effektstärken auf und ein nicht zu vernachlässigender Teil der Untersuchungen stellte sogar negative Effekte von Feedback in Lehr-/Lernsituationen fest.

### 3.2.4 Bedingungsfaktoren für die Effekte von Feedback

Wie sich in den vorangegangenen Abschnitten abgezeichnet hat, ist die Befundlage zu den Effekten von Feedback keinesfalls so konsistent wie noch von Ammons (1956) dargestellt. Einige Forscher begaben sich daher auf die Suche nach potenziellen Einflussfaktoren und Ursachen für die unterschiedlichen Ergebnisse von Feedback-Untersuchungen. Im Folgenden werden die verschiedenen potenziellen Bedingungsfaktoren, die in wissenschaftlichen Untersuchungen analysiert wurden, systematisiert und gebündelt vorgestellt.

#### Feedbackmerkmale

Wie bereits im Modell von Baldwin und Ford (1988) herausgearbeitet und in vielen Studien bestätigt wurde (siehe Kapitel 2.2.2), hat die methodische und inhaltliche Gestaltung eines Trainings wesentlichen Einfluss auf dessen Effektivität. Zentrale Feedback-Charakteristika und deren Wirkung sollen daher an dieser Stelle beschrieben werden.

*Positives versus Negatives Feedback:* Kluger und DeNisi (1996) haben in ihrer Meta-Analyse untersucht, inwiefern es einen Unterschied zwischen positivem und negativem Feedback gibt. Sie bewerteten die Feedback-Interventionen dafür auf einer sieben-stufigen Skala von „1 = sehr negativ“ bis „7 = sehr positiv“ für den Feedbackempfänger. Ob ein



Feedback positiv oder negativ war, hatte keine moderierende Wirkung auf den Effekt von Feedback.

*Komplexität des Feedbacks:* In einigen Studien wurden einfache Feedbackformen mit elaborierten Feedbackformen verglichen. Kulhavy, White, Topp, Chan und Adams (1985) führten eine Untersuchung zur Korrektoreffizienz von vier unterschiedlichen Feedbackarten beim Lernen aus Texten durch. Nach jedem der 16 Textabschnitte mussten die College-Studenten eine Multiple-Choice Aufgabe mit fünf Antwortalternativen lösen. Nach der Bearbeitung jeder Aufgabe erhielten alle Studenten ein Feedback. Dabei wurden vier unterschiedlich komplexe Feedbackarten additiv kombiniert: Korrekte Antwort (KCR), Falsche Antworten (KIR), Erklärung zu jeder falschen Antwortalternative (EIR), Relevante Textpassage zu korrekter Antwort (TEXT). Die Studenten mit einfachem Feedback korrigierten Fehler genauso gut wie die mit komplexem Feedback. Im Gegensatz dazu gibt es auch eine Reihe von Studien, die zeigen, dass elaboriertes Feedback lernförderlicher sein kann als einfache Feedbackformen (z.B. Winne, Graham & Prock, 1993). Butler, Godbole und Marsh (2013) führen als mögliche Ursache für die unterschiedlichen Befunde bezüglich der Effektivität verschiedener Feedbackformen an, dass es darauf ankommt, wie Lernen gemessen wird. In einem Experiment mit 60 Studenten verglichen sie drei Bedingungen miteinander: „knowledge of correct result“ (KCR), erklärendes Feedback, kein Feedback. Die Studenten mussten zunächst einen Text lesen und im Anschluss einen kurzen Test bearbeiten. Nach jeder Antwort erhielten die Studenten die korrekte Lösung, die korrekte Lösung sowie eine Erklärung dazu oder kein Feedback. Zwei Tage später mussten sie erneut den gleichen Test bearbeiten und erhielten zusätzliche Transfer-Fragen. Die Auswertung ergab, dass beide Feedbackgruppen eine bessere Leistung im Posttest zeigten als die Gruppe ohne Feedback. In den Posttest-Fragen, welche vom ersten Test wiederholt wurden, konnte kein Unterschied zwischen dem einfachen und dem elaborierten Feedback gefunden werden. Bei den Transfer-Aufgaben hingegen erreichte die Gruppe mit erklärendem Feedback eine signifikant bessere Leistung als die beiden anderen Gruppen. Dieser Befund deutet demnach darauf hin, dass ein elaboriertes Feedback hilfreicher ist, wenn nicht nur das Behalten von korrekten Lösungen, sondern das Verständnis des Lernmaterials erzielt werden soll. Eine weitere mögliche Einflussvariable könnte das Vorwissen sein. Beispielsweise fanden Winne et al. (1993) in ihrer Studie mit Schülern mit Lernschwierigkeiten einen Vorteil des ausführlicheren Feedbacks gegenüber dem einfachen Feedback (KR und KCR). Narciss (2006) fasst diese Ergebnisse mit der Erklärung zusammen, dass Lernende mit geringerem Vorwissen und Lernende, die komplexere Aufgaben bearbeiten müssen, von den zusätzlichen Informationen bei einem elaborierten Feedback im Vergleich zu einem einfachen Feedback profitieren. Auch Bangert-Drowns et al. (1991) kommen in ihrer Meta-Analyse zu

dem Fazit, dass Feedback, was nur Informationen über „richtig/falsch“ liefert, niedrigere Effekte hat als Feedback, was mehr Informationen bietet. Dabei sollte jedoch berücksichtigt werden, dass die Informationsmenge nicht die Aufmerksamkeits- und Verarbeitungskapazitäten des Feedbackempfängers überschreitet (Shute, 2008).

*Bezugsnorm:* Innerhalb eines Feedbacks kann der Vergleich mit einem sachlich begründeten Lernkriterium (kriteriumsorientiertes oder aufgabenbezogenes Feedback), mit der vorangegangenen Leistung (lernprozessorientiertes Feedback mit individueller Bezugsnorm) oder mit der Leistung anderer (wettbewerbsorientiertes Feedback mit sozialer Bezugsnorm) angestellt werden. In einer Untersuchung von Wollenschläger, Möller und Harms (2011) zeigten sich lernförderliche Effekte eines kriteriumsorientierten Feedbacks bei der Bearbeitung von Aufgaben zum wissenschaftlichen Denken. Krampen (1987) hat mit 385 Schülern eine Studie im Mathematikunterricht durchgeführt, bei der er die drei Feedbacktypen miteinander verglichen hat. Es stellte sich heraus, dass Feedback mit einer individuellen oder sachlichen Bezugsnorm motivationsfördernder ist als Feedback mit einer sozialen Bezugsnorm. Kluger und DeNisi (1996) konnten in ihrer Meta-Analyse dieses Ergebnis bezogen auf die Lernleistung nicht replizieren. Ob normative Informationen (Vergleich mit anderen) im Feedback enthalten waren, hatte keinen moderierenden Einfluss auf die Feedbackeffekte.

*Lob:* Sowohl in der Meta-Analyse von Kluger und DeNisi (1996) als auch in der von Hattie und Timperley (2007) hatten die Studien, welche Feedback in Form von Lob einsetzten kleinere Effektstärken im Vergleich zu anderen Studien.

*Extrinsische Belohnung:* Mithilfe des Datensatzes von Hattie (1999) haben Hattie und Timperley (2007) noch einmal 74 Meta-Analysen zusammenfassend betrachtet, die zumindest geringfügige Informationen über Feedback enthielten. Dabei stellten sie fest, dass unter anderem extrinsische Belohnungen am wenigsten effektiv waren, um Leistung zu steigern. Dieser Befund zeigte sich auch in Bezug auf die Motivation von Lernenden (Cameron & Pierce, 1994; Deci et al., 1999). In beiden Meta-Analysen hatten extrinsische Belohnungen in Form von Aufklebern, Auszeichnungen o.ä. einen signifikant negativen Effekt auf die intrinsische Motivation. Deci et al. (1999) bezeichnen diese eher als „contingencies“ und weniger als Feedback, da sie kaum Informationen zur Aufgabe enthalten.

*Intensität und Timing:* Bei Kluger und DeNisi (1996) hing die Höhe der Feedback-Effekte von der Anzahl der Feedbackepisoden pro Effekt ab. Es scheint demnach nicht nur die Qualität einer Feedbacknachricht, sondern auch die Quantität einen bedeutenden Einfluss auf deren Effekt zu haben. Welche Wirkung der Zeitpunkt der Feedbackpräsentation hat, haben Kulik und Kulik (1988) in ihrer Meta-Analyse von 53 Studien untersucht. In angewandten Studien mit realen Unterrichtsmaterialien hatte „immediate feedback“ signifikant

höhere Effekte auf das Lernen als „delayed feedback“. Im Kontext von experimentellen Laborsettings, bei denen es um das Lernen von Testmaterial ging, zeigte sich zum Teil ein gegenteiliges Ergebnis – hier hatten verzögerte Feedbackformen einen Vorteil gegenüber einem sofortigem Feedback. In der Literatur wird als Erklärung dafür die in Kapitel 3.2.2 beschriebene „interference-perseveration hypothesis“ von Kulhavy und Anderson (1972) angeführt. Hattie und Wollenschläger (2014) verweisen dementsprechend auf Ergebnisse, nach denen sofortiges Feedback effektiver ist, wenn prozedurales Wissen erlernt wird und weniger Transfer erforderlich ist. Im Gegensatz dazu eignet sich ihrer Ansicht nach verzögertes Feedback eher beim Erlernen von konzeptuellem Wissen und Aufgaben, bei denen ein hoher Transfer notwendig ist. Bangert-Drowns et al. (1991) haben die Effekte von Studien mit Feedback direkt nach jeder Antwort, direkt nach dem Test oder mit Verzögerung verglichen. Das sofortige Feedback hatte signifikant niedrigere Effekte. Dies konnten sie jedoch auf den Aufgabentyp zurückführen. Das sofortige Feedback wurde meist in Untersuchungen mit „programmed instruction“ eingesetzt, welche sich als weniger lernförderlich erwiesen als andere Lernformate.

### Aufgabenmerkmale

*Bekanntheitsgrad und Komplexität der Aufgaben:* Kluger und DeNisi (1996) konnten keinen signifikanten Zusammenhang zwischen der Neuheit der Aufgabe für die Lernenden und der Wirkung von Feedback finden. Es zeigte sich jedoch ein Interaktionseffekt mit der Dauer der Aufgabe. Bei unbekannten Aufgaben ist der Effekt größer, wenn die Aufgabe länger dauert - bei familiären Aufgaben gab es keine Interaktion mit der Aufgabendauer. Tendenziell geringere Effekte von Feedback stellten die Autoren bei einer hohen Komplexität der Aufgabe fest.

*Aufgabentyp:* Als einer von vier robusten Moderatoren in der Meta-Analyse von Kluger und DeNisi (1996) konnte der Aufgabentyp identifiziert werden. Bei Tests mit physischen Aufgaben und Tests, bei denen die Lernenden Regeln befolgen sollten, waren die Effekte von Feedback deutlich kleiner als bei anderen Aufgabentypen. In Bezug auf Gedächtnisaufgaben scheint Feedback hingegen eine größere Wirkung zu entfalten.

*Instruktionsansatz:* Die Unterschiede in den Effektstärken zwischen den zahlreichen Feedbackstudien konnten zudem durch den verwendeten Instruktionsansatz erklärt werden. Sowohl in der Meta-Analyse von Bangert-Drowns et al. (1991) als auch in der Synthese von Hattie (1999) hatten Feedbackinterventionen im Rahmen von Unterrichtseinheiten mit „programmed instruction“ eine geringere Wirkung auf das Lernen.

### Situative und methodische Aspekte

*Präsentation des Feedbacks:* Kluger und DeNisi (1996) fanden heraus, dass die Interventionen, bei denen Feedback mündlich kommuniziert wurde, signifikant kleinere Effekte hatten als andere Studien. Zur Frage, ob computer-unterstützte Feedbackinterventionen effektiver oder weniger effektiv sind, gibt es unterschiedliche Befunde. In der Meta-Analyse von Kluger und DeNisi (1996) hatte computer-unterstütztes Feedback tendenziell geringere Effekte. Bei Hattie (1999) zeigt sich ein gegenteiliges Bild. Hier hatte computer-unterstütztes Feedback mit die größten Effektstärken. Auch Video- und Audio-Feedback hatte verglichen mit anderen Methoden hohe Effekte. Unabhängig vom Präsentationsformat sollte sichergestellt sein, dass das Feedback klar und verständlich formuliert wird (Kopp & Mandl, 2014).

*„Control for presearch availability“:* Die zu frühe Verfügbarkeit von Feedback-Informationen wurde von Kulhavy (1977) als „presearch availability“ bezeichnet und wird als eine wesentliche Ursache von negativen Feedbackeffekten betrachtet. Wenn Lernende zu früh ein Feedback über die korrekte Lösung erhalten, müssen sie die Antworten nur abschreiben, ohne vorher das Lernmaterial richtig gelesen und eine eigene Lösung erarbeitet zu haben. Dies könnte auch eine mögliche Erklärung für die niedrigen Effekte von „programmed instruction“-Ansätzen sein. Durch die kleinschrittige Präsentation der Lerninhalte bei programmierten Unterrichtseinheiten sind Aufgaben und Feedbackinformationen häufig auf einer Seite ersichtlich, was zu einer Art Redundanz und somit zu fehlenden Feedbackeffekten führt. Die Ergebnisse von Bangert-Drowns et al. (1991) belegen dies. Wenn Schüler Zugang zu den Feedbackinhalten hatten, bevor sie selbst eine Antwort formuliert haben, ging dies mit signifikant geringeren Effektstärken einher.

*Zielsetzungsmaßnahmen:* In der zusammenfassenden Betrachtung von 74 Meta-Analysen von Hattie (1999) zählten die Feedbackformen, die sich auf bestimmte Ziele beziehen, zu den effektivsten. In ihrer „goal setting theory“ haben Locke und Latham (1990) den Zusammenhang zwischen Zielen und Performanz thematisiert. Der Kern dieser Theorie ist die Annahme, dass spezifische höhere Ziele zu einer höheren Leistung führen als eher moderate Ziele. Diese Behauptung basiert auf Befunden, die Locke, Shaw, Saari und Latham (1981) in einem Review zusammengetragen haben. Sie fanden heraus, dass in über 90 % der Studien spezifische, herausfordernde Ziele zu einer höheren Leistung führten als einfache „do your best“-Ziele. Diese Hypothese entspricht einem positiven linearen Zusammenhang und widerspricht auf den ersten Blick dem „Yerkes-Dodson-Gesetz“, welches aussagt, dass Leistung bei einem moderaten Level von Motivation am höchsten ist. Locke et al. (1981) halten diesbezüglich fest, dass es sich nicht wirklich um einen Widerspruch handelt. Sie gehen auch davon aus, dass Ziele durchaus aufgegeben werden können, wenn sie zu schwer erreichbar sind. Allerdings betonen sie dabei, dass ihre aufgestellte Hypothese eine gewisse

vorherige Verpflichtung oder Bindung gegenüber dem Ziel unterstellt. Birney, Burdick und Teevan (1969) haben zusammenfassend angemerkt, dass Ziele für die individuelle Person nicht zu hoch, aber auch nicht zu niedrig angesetzt werden sollten. Wenn Ziele so hoch gesteckt sind, dass sie unerreichbar sind, ist die Wahrscheinlichkeit für Misserfolg sehr hoch. Dies führt wiederum zu Entmutigung und niedriger Motivation. Wenn im Gegenzug Ziele so niedrig angesetzt werden, dass ihr Erreichen gewiss ist, verlieren sie ebenfalls ihre motivierende Funktion und ihr Potenzial, Anstrengung zu fördern. Diese Annahme findet sich auch in der Selbstbestimmungstheorie von Deci und Ryan (1993) wieder. Demnach wird intrinsische Motivation nur erzeugt, wenn ein optimales Anforderungsniveau besteht. Dies ist dann der Fall, wenn zwischen den Anforderungen eines Ziels und den eigenen Fähigkeiten ein optimales Verhältnis herrscht, das angestrebte Ziel also weder als zu leicht noch zu schwer empfunden wird. Kluger und DeNisi (1996) unterschieden schwierige Ziele sowie moderate, leichte und keine Ziele. Es zeigte sich ein tendenziell signifikanter Effekt ( $0.01 < p < 0.05$ ) von Zielsetzungsmaßnahmen mit höheren Effekten bei höheren Zielen. Die Tatsache, dass dieser Effekt nicht statistisch belastbar nachgewiesen werden konnte, führen die Autoren darauf zurück, dass in ihrer Meta-Analyse nur ein geringer Anteil der Studien Ziele manipuliert hat. Malone (1981) betont zudem, dass Ziele bestimmte Voraussetzungen erfüllen müssen, um den Lerner intrinsisch zu motivieren. Wichtig ist zum Beispiel, dass die Ziele für den Lernenden persönlich von Bedeutung sind. Auch Locke et al. (1981) gehen davon aus, dass Lernende Ziele bewusst akzeptieren müssen, wenn diese die Leistung steigern sollen.

### Personale Faktoren

Für die Interpretation der Wirkung personaler Variablen bei der Übermittlung von Feedback können allgemeine Sender-Nachricht-Empfänger-Modelle aus der Kommunikationsforschung herangezogen werden. Diese drücken im Grunde aus, dass eine Person (Sender) einer anderen Person (Empfänger) eine Nachricht mit bestimmten Intentionen sendet, welche schließlich vom Empfänger dekodiert wird. Die gesendete Feedback-Botschaft muss demnach nicht mit der empfangenen Feedback-Botschaft übereinstimmen. Nach Schulz von Thun (2013) beinhaltet eine Nachricht immer Botschaften auf vier Ebenen: Sachinhalt, Selbstoffenbarung, Beziehung und Appell. Je nachdem wie die vier Seiten einer Nachricht von Sender und Empfänger unterschiedlich gewichtet werden, kann es zu Missverständnissen kommen. Strijbos und Müller (2014) ergänzen diesen Ansatz in ihrem Rahmenmodell zur Gestaltung und Verarbeitung von Feedback um situative und interaktionale Elemente. Personale Variablen, Rahmenbedingungen und die Repräsentation des Gegenübers beeinflussen die Gestaltung und die Verarbeitung von Feedback und können dabei auch miteinander interagieren. Zum Beispiel kann ein Feedbackgeber negatives Feedback vermei-

den, um den anderen nicht zu verletzen oder negatives Feedback verzerren, wenn der Empfänger als verletzlich wahrgenommen wird. Unter situationalen Rahmenbedingungen bei der Gestaltung von Feedback verstehen die Autoren Aspekte wie den Zeitpunkt oder das Setting des Feedbacks, welche bereits ausführlich thematisiert wurden. Zur Frage, welche Rolle Merkmale des Feedbackgebers (z.B. domänenspezifisches Wissen oder die Fähigkeit, Feedback zu geben) für die Gestaltung und Wirkung von Feedback hat, gibt es bisher kaum belastbare Forschungsergebnisse. Einige Studien liefern Hinweise dafür, dass die wahrgenommene Kompetenz des Senders bedeutend für die Zufriedenheit mit dem Feedback (z.B. Klein, Kraut & Wolfson, 1971) und der Reaktion auf Feedback ist (z.B. Raemdonck & Strijbos, 2013). Die Verarbeitung von Feedback unter dem Einfluss personaler Variablen des Feedbackempfängers wurde wesentlich häufiger untersucht und soll im Folgenden thematisiert werden.

Drei personale Merkmale des Feedbackempfängers wurden in den letzten Jahren intensiv erforscht: Attribution, Selbstwirksamkeit und Zielorientierung.

*Attribution und Selbstwirksamkeit:* Wie bereits beschrieben, können die Ursachen eines Handlungsergebnisses und eines dementsprechenden Feedbacks individuell unterschiedlich bewertet werden. In Kapitel 3.2.2 wurden verschiedene Belege dafür aufgeführt, dass Feedback mithilfe von attributionalen Hinweisen die Ursachenzuschreibung und somit die Leistung von Lernenden steigern kann. Andersherum betrachtet, kann die Neigung von Personen, Handlungen auf bestimmte Ursachen zu attribuieren, die Wirkung von Feedback beeinflussen. Bei einer internal, variablen und kontrollierbaren Ursachenzuschreibung sind demnach eher positive Auswirkungen nach einem Feedback zu erwarten. Die Attribution der Feedbackinhalte steht außerdem in einem wechselseitigen Zusammenhang zur Selbstwirksamkeitserwartung des Feedbackempfängers. Es kann davon ausgegangen werden, dass ein Feedback nur dann positive Selbstwirksamkeitserwartungen erzeugt, wenn die vorherige Leistung als beeinflussbar wahrgenommen wird. Die Ursachenzuschreibung von Personen in bestimmten Situationen basieren wiederum auf den Selbstwirksamkeitserwartungen, die sie aufgrund vorheriger Leistungen entwickelt haben. Das Zusammenspiel dieser verschiedenen Faktoren im Rahmen von Feedbackinterventionen wurde in der Studie von Lyden, Chaney, Danehower und Houston (2002) belegt. Der direkte und implizite leistungssteigernde Effekt von positiven Selbstwirksamkeitserwartungen wurde vielfach empirisch nachgewiesen (z.B. Heslin & Latham, 2004; Karl, O'Leary-Kelly & Martocchio, 1993).

*Zielorientierung:* Im Zusammenhang mit dem Einfluss von Zielen auf die Effekte von Feedback spielen auch Überzeugungen bei der Zielerreichung eine wesentliche Rolle. Diese werden in der Literatur als Zielorientierungen bezeichnet. Dabei werden klassischerweise zwei Dimensionen differenziert: Performanz- oder Leistungsorientierung und Lernorientie-

rung (Dweck, 1986). Lernorientierung ist charakterisiert durch den Drang, die eigenen Kompetenzen zu verbessern. Personen mit einer Performanzorientierung hingegen geht es eher darum, gegenüber anderen die eigenen Fähigkeiten zu präsentieren und von diesen als kompetent eingeschätzt zu werden. Hoska (1993) unterscheidet zudem zwischen „acquisition goals“ und „avoidance goals“. Während „acquisition goals“ dem Lernenden dabei helfen, sich etwas Wünschenswertes anzueignen, sind „avoidance goals“ eher darauf ausgerichtet, etwas Unerwünschtes zu vermeiden. Mithilfe einer Meta-Analyse haben Payne, Youngcourt und Beaubien (2007) 160 Studien hinsichtlich der Wirkung der Zielorientierung in verschiedenen Lernkontexten überprüft. Die Autoren fanden einen positiven Zusammenhang zwischen Lernorientierung und Leistung, aber keinen Zusammenhang zur Performanzorientierung. Lernorientierung scheint sich demnach deutlich günstiger auf den Leistungserwerb auszuwirken. Dieser Effekt konnte im Großen und Ganzen auch im Kontext von Feedbackinterventionen bestätigt werden (z.B. Dahling & Ruppel, 2016; Heslin & Latham, 2004). Allerdings weisen die Ergebnisse einiger Studien darauf hin, dass es Unterschiede in den Effekten zwischen positivem und negativem Feedback gibt. Bei negativem Feedback scheint eine Lernorientierung dazu beizutragen, selbstwertgefährdende Effekte abzumildern, wohingegen eine Performanzorientierung zu Frustration und Anstrengungsreduktion führt (z.B. Elliott & Dweck, 1988). Cianci, Schaubroeck und McGill (2010) haben in einer experimentellen Studie mit einem 3 x 2-faktoriellen Design 113 Studenten randomisiert verschiedenen Bedingungen zugewiesen. Eine Hälfte der Studenten erhielt positives Feedback, die andere negatives Feedback. Außerdem gab es drei verschiedene Zielbedingungen: Instruktion mit Lernorientierung, Leistungsorientierung oder keinen Zielen. Die Gruppe mit Lernorientierungs-Instruktionen zeigte nach negativem Feedback eine höhere Leistung, aber nach positivem Feedback eine schlechtere Leistung. Eine mögliche Erklärung dafür könnte sein, dass bei sehr positivem Feedback das befriedigte Kompetenzerleben bei performanzorientierten Lernenden zu einer Leistungssteigerung führt, während Personen mit einer hohen Lernorientierung das Gefühl haben, nicht viel mehr lernen zu können.

*Persönlichkeit:* In wenigen Studien wurde die Rolle der Persönlichkeitseigenschaften Neurotizismus, Extraversion, Offenheit, Verträglichkeit und Gewissenhaftigkeit für die Verarbeitung von Feedback untersucht. Einzelne Befunde deuten auf eine Abhängigkeit der Feedback-Nutzung von der Persönlichkeit des Feedbackempfängers hin. Bell und Arthur (2008) stellten beispielsweise einen signifikanten Zusammenhang zwischen der Verträglichkeit der Versuchspersonen und der Akzeptanz von Feedback fest. Insgesamt scheint der Einfluss von Persönlichkeit auf die Feedbackverarbeitung jedoch eher gering zu sein (Atwater & Brett, 2005).

Die Motivation des Feedbackempfängers war bisher kaum Gegenstand von Untersuchungen zu Feedbackeffekten. In der allgemeinen Lehr-/Lernforschung gibt es jedoch zahlreiche empirische Belege dafür, dass eine höhere (insbesondere intrinsische) Motivation beim Lernenden mit einer höheren Lernleistung einhergeht (Schiefele & Schreyer, 1994).

### 3.2.5 Theoretische Erklärungsansätze für die Wirkung von Feedback

Zur Darstellung der vielfältigen Wirkungsweisen von Feedback und deren Bedingungen gibt es verschiedene theoretische Modelle, welche im Folgenden vorgestellt werden sollen.

#### Das „Response-Certitude-Model“ von Kulhavy und Stock (1989)

Innerhalb der kognitiven Feedbackforschung hat das kybernetische Feedbackmodell von Kulhavy und Stock (1989) große Aufmerksamkeit erhalten. Im Mittelpunkt des Modells stehen kognitive Prozesse bei der Verarbeitung von Feedback. In Bezug auf das Feedback unterscheiden die Autoren zwischen zwei inhaltlichen Komponenten: Während die Verifikationskomponente den Lernenden über die Korrektheit einer Antwort informiert, liefert die Elaborationskomponente zusätzliche Informationen zur Lösung der Aufgabe. Zur Vorhersage der Wirkung von Feedback bedienen sie sich dem systemtheoretischen Konzept von Feedback. Dieses besagt, dass die Diskrepanz zwischen einem Ist- und einem Soll-Wert regulierende Handlungen auslöst, welche diese reduzieren sollen. Die Stärke der Regulierungsaktivitäten hängt dabei vom Ausmaß der Diskrepanz ab (vgl. Kapitel 3.2.1). Die Höhe der Diskrepanz wird nach Kulhavy und Stock (1989) wiederum davon bestimmt, wie sicher sich die Lernenden ihrer Antwort sind. Diese Antwortsicherheit wird als Maß für das Verständnis der Lernenden betrachtet und in der Regel mithilfe einer fünf-stufigen Likert-Skala durch Selbsteinschätzung erfasst. Wenn ein Lernender beispielsweise überzeugt davon ist, die richtige Antwort gegeben zu haben (hohe Antwortsicherheit), dann jedoch das Feedback bekommt, dass die Antwort falsch ist, führt dies zu einer hohen Diskrepanz. In der Folge müsste dies den Lernenden dazu anregen, sich intensiv mit den Informationen auseinanderzusetzen, um die Diskrepanz zu verringern. Die Wahrscheinlichkeit von korrekten Antworten im Posttest wird daher von der Anstrengung des Lernalters bei der Feedbackverarbeitung (operationalisiert durch Bearbeitungszeit) beeinflusst.

Verschiedene Studien zum Lernen aus Texten bestätigen die Annahme dieses Modells, dass die Feedbackverarbeitung von der Diskrepanz zwischen der eigenen Antwort und der zurückgemeldeten Lösung beeinflusst wird und diese mit der Antwortsicherheit der Lernenden zusammenhängt (Mory, 2004). Die Antwortsicherheit wird jedoch über eine Selbst-



einschätzung der Lernenden erfasst und es stellt sich daraus die Frage, ob dies als reliables und valides Maß für Verständnis betrachtet werden kann.

### Das „Mindful-Processing-Model“ von Bangert-Drowns et al. (1991)

Basierend auf vorhandenen Reviews sowie der oben beschriebenen Meta-Analyse entwickelten Bangert-Drowns et al. (1991) ein fünf-stufiges Modell zur Feedbackverarbeitung. Es handelt sich dabei um ein recht allgemeines Modell, welches die Bedeutung von motivationalen und kognitiven Faktoren im Feedbackverarbeitungsprozess betont. Im Mittelpunkt des Modells steht das von Salomon und Globerson (1987) eingeführte Konstrukt des „mindful processing“ von Informationen. Sie definieren „mindfulness“ als „*volitional, meta-cognitively guided employment of non-automatic, usually effortful processes*“ (Salomon & Globerson, 1987, S. 623). Das Vorhandensein dieser zielgerichteten, aktiven und intensiven Verarbeitung von Informationen betrachten sie als einen wesentlichen Einflussfaktor dafür, ob Lernende tatsächlich das umsetzen, was sie können. Bangert-Drowns et al. (1991) ziehen daraus die Schlussfolgerung, dass Feedback vor allem dann effektiv ist, wenn Lernende zum „mindful processing“ angeregt werden. Die Verarbeitung von Feedback wird durch das Interesse, die Zielorientierung, die Selbstwirksamkeit, das relevante Vorwissen und die Antwortsicherheit der Person bestimmt.

In Anlehnung an Kulhavy (1977) betrachten Bangert-Drowns et al. (1991) das Fehlen von „presearch availability“ als eine wichtige Voraussetzung für das Zustandekommen der bewussten Verarbeitung in den genannten Stufen. Wenn das Feedback bereits verfügbar wäre, bevor der Lernende selbst im Gedächtnis nach der richtigen Antwort sucht, wird die Feedbackverarbeitung abgebrochen und es kommt nicht zu einem „mindful processing“.

### Das „Feedback-Modell für selbstreguliertes Lernen“ von Butler und Winne (1995)

Butler und Winne (1995) systematisieren die Befunde der Feedbackforschung im Kontext von Erkenntnissen zum selbstregulierten Lernen. Sie betrachten die Wirkungen verschiedener Feedbackarten in Abhängigkeit von individuellen kognitiven und motivationalen Merkmalen der lernenden Person. Entsprechend dem Konzept des selbstregulierten Lernens hat das so genannte „monitoring“ des Lernprozesses einen großen Stellenwert in ihrem Modell. Auf der Grundlage des eigenen Wissens und der motivationalen Überzeugungen entwickelt ein Lernender beim selbstregulierten Lernen eine subjektive Repräsentation der Eigenschaften und Anforderungen einer Aufgabe. Basierend auf dieser Aufgabenrepräsentation setzt sich der Lernende Ziele, zu deren Erreichung verschiedene taktische Maßnahmen und Strategien angewendet werden. Im Zuge des „monitoring“ der ausgewählten Handlungsstrategien und der daraus resultierenden Produkte wird ein internes Feedback generiert. Das interne Feedback dient als Grundlage dafür, die Aufgabe und den ei-

genen Umgang damit zu überdenken. Es kann beispielsweise dazu führen, dass neue Ziele gesetzt oder bestehende Ziele modifiziert werden sowie die Strategiewahl überprüft und ggf. angepasst wird. Als Folge des „monitoring“-Prozesses werden ggf. Motivation und Wissen verändert, was wiederum folgende Selbstregulationsprozesse beeinflussen kann. Wenn der Lernende zusätzlich ein externes Feedback erhält, können die darin enthaltenen Informationen das interne Feedback bestätigen, ergänzen oder im Widerspruch dazu stehen. Die Wirkung des externen Feedbacks ist demnach abhängig von der Übereinstimmung zwischen internem und externem Feedback und davon, wie die lernende Person auf eine mögliche Diskrepanz reagiert. Abgeleitet von den Annahmen von Carver & Scheier, 1990s; Carver und Scheier (1990) zur affektbedingten Selbstregulation, stellen Butler und Winne (1995) die Hypothese auf, dass eine Diskrepanz zwischen internem und externem Feedback prinzipiell den Lernenden dazu anregt, Lösungen zu suchen, um diese Diskrepanz zu reduzieren. Hinsichtlich der Bedeutung unterschiedlicher Inhalte des externen Feedbacks leiten Butler und Winne (1995) bezugnehmend auf die Erkenntnisse von Balzer, Doherty und O'Connor (1989) zwei Implikationen ab:

1. Ergebnisorientierte Feedbackarten helfen dem Lernenden am wenigsten beim selbst-regulierten Lernen.
2. Elaborierte Feedbackformen hingegen können den Lernenden dabei unterstützen, das eigene Lernen zu steuern. Indem Lernenden dabei geholfen wird, wichtige Hinweise (z.B. Aufgabenmerkmale) und deren Bedeutung für den Lernfortschritt zu erkennen, wird ihnen das „monitoring“ erleichtert, welches essenziell für selbst-reguliertes Lernen ist.

Butler und Winne (1995) schlagen vor, dass elaboriertes Feedback zwei Komponenten beinhalten sollte. Während die produktorientierte Komponente inhaltliche Informationen in der entsprechenden Domäne liefern sollte, ist die prozessorientierte Komponente dafür verantwortlich, strategische Informationen für die Verarbeitung des domänenspezifischen Wissens bereitzustellen.

### Die „Feedback Intervention Theory“ von Kluger und DeNisi (1996)

Ausgangspunkt für die „Feedback Intervention Theory“ (FIT) war die in Kapitel 3.2.3 skizzierte Meta-Analyse von Kluger und DeNisi (1996). Die Autoren gehen davon aus, dass es vier verschiedene Strategien gibt, um die Diskrepanz zwischen Ist und Soll zu reduzieren: a) Die Anstrengung wird erhöht, wenn das Ziel klar ist, eine hohe Verbindlichkeit gegeben ist und der Glaube an den Erfolg groß ist, b) Das Ziel wird aufgegeben, wenn es als unwahrscheinlich eingeschätzt wird, dass die Diskrepanz durch Handlungen verringert werden kann, c) Wenn das Ziel nicht aufgegeben werden soll, kann es auch verändert werden – bei

negativem Feedback wird das Ziel „heruntergeschraubt“ und bei positivem Feedback wird das Ziel nach oben gesetzt, d) Die Diskrepanz kann letztlich auch durch die Ablehnung des Feedbacks überwunden werden.

Bezogen auf die Steuerung von Leistung werden drei Prozessebenen unterschieden. Die höchste Ebene der „meta-task processes“ ist vor allem dafür verantwortlich, selbstbezogene Ziele zu regulieren. In der Ebene der „task-motivation processes“ steht die aufgabenbezogene Motivation im Mittelpunkt. Bei der Ebene der „task-learning processes“ geht es darum, die einzelnen Schritte bei der Aufgabenbearbeitung zu steuern. Die Autoren gehen davon aus, dass Feedback in erster Linie auf der aufgabenbezogenen Motivationsebene verarbeitet wird. Wenn der Ist-Soll-Vergleich mithilfe des Feedbacks negativ ausfällt, wird die Anstrengung erhöht. Wenn die Diskrepanz auf diesem Weg nicht eliminiert wird, kann das Feedback die Aufmerksamkeit in Abhängigkeit von Aufgabenmerkmalen, der Art des Feedbacks sowie von situativen und personenspezifischen Faktoren auf die niedrigere oder höhere Ebene lenken. Durch lösungs- oder korrekturrelevante Hinweise im Feedback (z.B. „Du benutzt deinen Daumen beim Tippen nicht“) kann die Aufmerksamkeit auf die Ebene der „task-learning processes“ gerichtet werden. Kluger und DeNisi (1996) formulieren in ihrem Modell die Hypothese, dass die Effekte von Feedback durch Hinweise, welche die Aufmerksamkeit auf das Selbst („meta-task processes“) lenken, abgeschwächt werden und im Gegensatz dazu größer sind, wenn das Feedback Hinweise enthält, die die Aufmerksamkeit auf die Motivations- oder die Lern-Ebene lenken. Diese Annahme konnte durch die Ergebnisse ihrer Meta-Analyse bestätigt werden.

#### Das „Interactive-Two-Feedback-Loops-Model“ von Narciss (2006)

Narciss (2006) betrachtet Lernen als einen Prozess der aktiven Konstruktion und Kommunikation von Wissen. Das Feedback-Modell basiert auf der Vorstellung, dass Feedback keine reine Wissensvermittlung im Sinne von evaluativen, ergebnisorientierten Informationen darstellen sollte, sondern die Aufgabe hat, den Lernenden beim konstruktiven Prozess des Wissenserwerbs zu unterstützen. Sie betrachtet Feedback als eine Art Tutor, welcher Informationen für die Bewältigung der Aufgabe bereitstellt, ohne unmittelbar die Lösung zu geben. Die Grundlage des „Interactive-Two-Feedback-Loops-Model“ (ITFL) stellt ein allgemeiner Regelkreis dar, welcher aus einer Regelstrecke und einer Regeleinrichtung besteht. Unter einer Regelstrecke wird im ITFL-Modell die Bearbeitung von Lernaufgaben verstanden. Zu einer Regelstrecke gehören eine Regelgröße (legt fest, was reguliert werden soll, z.B. Kompetenz), ein Messfühler (misst die Ausprägung der Regelgröße) und ein Stellglied (leitet Korrekturmaßnahmen ein). Wesentliche Bestandteile einer Regeleinrichtung sind ein Soll-Wert-Einsteller (legt die Führungsgröße fest, also welchen Wert die Regelgröße haben

soll), ein Vergleichsmodul (vergleicht den Ist-Wert mit dem Soll-Wert) sowie ein Regler (leitet aus der Diskrepanz zwischen Ist- und Soll-Wert eine Stellgröße ab).

In Bezug auf eine Lernsituation mit Feedback gibt es zwei interagierende Regelkreise. Während im internen Regelkreis die interne Ist-Größe (z.B. Antwortsicherheit) verarbeitet wird, geht es im externen Regelkreis um die Verarbeitung der durch das Lehrmedium erfassten Ist-Größe. Für die Regulation der beiden Regelkreise wird in einem ersten Schritt ein Soll-Wert bestimmt. Der interne Soll-Wert basiert auf einer subjektiven Repräsentation der Aufgaben-Anforderungen, welche wiederum von individuellen kognitiven, meta-kognitiven und motivationalen Voraussetzungen der lernenden Person beeinflusst wird. Der externe Soll-Wert wird auf der Basis externer Aufgaben-Anforderungen generiert. Diese ergeben sich durch situative Merkmale der Lernsituation (z.B. Lerninhalte, Lernziele). Mithilfe eines internen und eines externen Messfühlers wird von der lernenden Person selbst ein interner Ist-Wert und von der Feedbackquelle ein externer Ist-Wert gemessen. Innerhalb der beiden Regeleinrichtungen werden die internen bzw. die externen Ist- und Soll-Werte verarbeitet und jeweils miteinander verglichen. Im externen Regelkreis führt das Ergebnis des Vergleichs zwischen externem Ist- und Soll-Wert zu der Bestimmung einer Stellgröße (z.B. Fehler zurückmelden), welche in Form eines externen Feedbacks an die interne Regeleinrichtung übermittelt wird. Zusätzlich zu Verarbeitung und Vergleich des internen Ist- und Soll-Werts findet im internen Regelkreis daher ein Vergleich zwischen dem externen Feedback und dem internen Ist-Wert sowie dem internen Soll-Wert statt. Basierend auf diesen Vergleichen entwickelt die lernende Person eine interne Stellgröße, welche das Fundament für korrigierende Maßnahmen des Stellglieds ist (z.B. Anstrengen). Diese Korrektur-Prozesse haben schließlich Auswirkungen auf die Regelgröße, deren Ausprägung (Ist-Wert) dann wieder durch den internen und externen Messfühler an die Regeleinrichtungen zurückgemeldet werden.

Auf der Grundlage ihres Modells führte Narciss (2006) verschiedene Studien mit Schülern und Studenten zu den Effekten von informativen tutoriellen Feedbackformen durch. In allen Studien führte ITF zu besseren Lernergebnissen als einfache ergebnisorientierte Feedbackarten in Form von „knowledge of result“ (KR).

### Das „visible learning and teaching model“ von Hattie und Timperley (2007)

Basierend auf dem „visible learning and teaching model“ (Hattie & Timperley, 2007; Hattie & Gan, 2011) nehmen Hattie und Wollenschläger (2014) mit Bezug auf empirische Erkenntnisse der Pädagogischen Psychologie eine multidimensionale Konzeptualisierung der Wirkungsweise von Feedback vor. Nach der Auffassung von Hattie und Timperley (2007) ist das oberste Ziel von Feedback die Diskrepanz zwischen der Leistung einer lernenden Person und bestimmten Zielen zu reduzieren. Damit dies gelingen kann, muss Feedback ihrer

Ansicht nach vor allem drei Fragen beantworten: „Where am I going?“, „How am I going?“ und „Where to next?“. Sie postulieren, dass Lernende keinen Bedarf sehen, die Lücke zwischen Ist und Soll zu schließen, solange ihnen das Ziel nicht klar ist. Feedback sollte dementsprechend hinsichtlich der ersten Frage Informationen zu den Zielen und Erfolgskriterien der Aufgabe liefern. Das eigentliche Feedback wird über die Beantwortung der zweiten Frage bereitgestellt. Der Lernende sollte demnach über die aktuelle Leistung in Relation zu einem erwarteten Standard informiert werden. Die Autoren gehen davon aus, dass es nicht ausreicht, wenn Feedback nur Auskunft über die erreichte Leistung gibt und betonen die Bedeutung der dritten Frage „Where to next?“. Ein lernförderliches Feedback muss dem Lernenden daher Hinweise darauf geben, wie die Leistung verbessert werden kann.

Wie effektiv die Antworten auf diese Fragen dabei helfen, die Lücke zwischen Ist und Soll zu schließen, hängt davon ab, auf welcher Ebene das Feedback gegeben wird. Hattie und Timperley (2007) unterscheiden vier Feedback-Ebenen: „task level“, „process level“, „self-regulation level“ und „self level“. Da ein Feedback bezüglich der Person (z.B. Lob in Form von „Du bist klug“) nicht in der Lage ist, Antworten auf die oben genannten Fragen zu geben, schließen es Hattie und Wollenschläger (2014) aus ihrem Modell aus. Sie gehen deshalb von drei Ebenen aus, auf denen Feedback wirken kann. Feedback auf dem „task level“ teilt der lernenden Person mit, wie gut die Aufgabe gelöst wurde. Wird auf dieser Ebene dem Lernenden die richtige Antwort gegeben (KCR), würde dies die Frage „Where am I going?“ beantworten. Ein Feedback auf dem „task level“ in Form eines „knowledge of result“ oder „knowledge of performance“ geht auf die Frage „How am I going?“ ein. Wird dies zum Beispiel mit Hinweisen zu den Fehlern verknüpft, handelt es sich um ein Feedback in Bezug auf die Frage „Where to next?“. Auf dem „process level“ bezieht sich das Feedback auf die Prozesse, die einer Aufgabe und deren Bearbeitung zugrunde liegen (z.B. Strategien zur besseren Informationssuche oder Fehlererkennung). Wenn der Lerner auf dieser Ebene zusätzliche Informationen zu den Zielen des Lernmaterials bekommt, kann dies die erste der drei Fragen beantworten. Lernende bekommen auf dem „process level“ Informationen zu ihrer eigenen Leistung, wenn sie auf ihre individuellen Fehler hingewiesen werden und sie Auskunft darüber erhalten, warum ihre Antwort richtig oder falsch war. Konkrete Hinweise für die richtige Aufgabenbearbeitung können dem Lernenden dabei helfen, herauszufinden, wie die eigene Leistung verbessert werden kann („Where to next?“). In Bezug auf die „meta-task“-Ebene können Hinweise zu meta-kognitiven Strategien (z.B. Selbsteinschätzung, Überwachen des Lernprozesses oder Attributionsverhalten) alle drei Fragen beantworten.

Wie bereits erwähnt, hängt die Effektivität von Feedback von der Ebene ab, auf der es übermittelt wird. Welche der drei Ebenen zu bevorzugen ist, hängt wiederum von ver-

schiedenen Faktoren ab. Nach Hattie und Wollenschläger (2014) ist ein einfaches ergebnisorientiertes Feedback auf dem „task level“ am sinnvollsten bei Novizen und Aufgaben, die eine geringe Schwierigkeit aufweisen und wenig Transfer erfordern. Wenn es sich um kompetentere Lerner oder Aufgaben mit einer hohen Komplexität und großem Ausmaß an Transfer handelt, eignet sich eher ein Feedback auf dem „process level“. Des Weiteren gehen die Autoren davon aus, dass auf der „task“-Ebene ein sofortiges Feedback effektiver und auf dem „process level“ ein verzögertes Feedback sinnvoller ist (vgl. Kapitel 3.2.4). In Hinsicht auf die „meta-task“-Ebene betonen Hattie und Wollenschläger (2014) die Bedeutung von Eigenschaften der Person, die einen Einfluss auf die Feedbackverarbeitung haben. Dazu zählen sie unter anderem das Vorwissen, die Ziele, die wahrgenommen Selbstwirksamkeit, das Attributionsverhalten oder die Antwortsicherheit einer Person (siehe Ergebnisse in Kapitel 3.2.4).

### 3.2.6 Feedback im (früh-) pädagogischen Kontext

Die Literatur zu Feedback bezieht sich, wie bereits erwähnt, häufig auf typische Lehr-/Lernsituationen im Kontext von Schule und Universität. Im Folgenden soll daher darauf eingegangen werden, inwiefern Feedback im Bereich der beruflichen Weiterbildung von Frühpädagogen eingesetzt wird bzw. werden kann.

Ein Format, welches mit video-basiertem Feedback arbeitet, ist die „Video Interaction Guidance“-Methode (VIG). VIG wurde ursprünglich in den Niederlanden zur Betreuung von Familien mit Erziehungsschwierigkeiten entwickelt. Mittlerweile wird es weltweit auch dazu eingesetzt, um die Interaktionsqualität zwischen pädagogischen Fachkräften und Kindern zu verbessern. Im Mittelpunkt des Trainings stehen das feinfühlige Verhalten und die emotionale Unterstützung der Erzieherin. Mithilfe von Videobildern wird das eigene Interaktionsverhalten reflektiert und analysiert. Dazu werden Videoaufnahmen im beruflichen Alltag der Fachkräfte angefertigt, welche vom Trainer in kurze Sequenzen unterteilt werden. In der darauffolgenden Besprechung werden diese ausführlich diskutiert. Im Sinne der positiven Verstärkung werden in erster Linie positive Verhaltensbeispiele gezeigt und besprochen. Fukkink und Tavecchio (2010) haben die Wirksamkeit von VIG in einer experimentellen Kontrollgruppen-Studie mit 95 Frühpädagogen überprüft. Das Training bestand aus vier Sitzungen, in denen jeweils kurze Videosequenzen besprochen wurden. Das Interaktionsverhalten wurde mit der CIS, einer eigens entwickelten Skala zur sensiblen Responsivität und einem Kodierschema zu VIG-relevanten Aspekten bewertet. Für die Einschätzung der einzelnen Variablen standen jeweils 10 bis 15 Minuten Videomaterial zur Verfügung. Die Analysen zeigten einen positiven Effekt des Trainings auf das sensitive und stimulierende Verhalten der Erzieherinnen. In Deutschland findet sich ein ähnliches Vorgehen bei der „Vi-

deo-Interaktions-Begleitung“ (VIB) oder „Marte Meo“ (Kieferle & Schauland, 2016). Kritisch anzumerken ist, dass in beiden Fällen die Analyse der Videoaufnahmen durch den Trainer, auf welcher das Reflexionsgespräch basiert, nicht standardisiert mithilfe eines erprobten Beobachtungsverfahrens durchgeführt wird und eine Qualitätssicherung daher nur schwer möglich ist. Empirisch belastbare Ergebnisse zu deren Wirksamkeit gibt es bislang nicht.

In einer Studie von Werner, Vermeer, Linting und van IJzendoorn (2018) werden die Effekte von Videofeedback zur Verbesserung der Feinfühligkeit von pädagogischen Fachkräften untersucht. Dafür wurde die „Video-feedback Intervention to promote Positive Parenting and Sensitive Discipline for Child Care (VIPPCC)“ auf das Kindergartensetting übertragen. Das Training besteht insgesamt aus fünf Sitzungen, in denen jeweils Videoaufnahmen aus dem Alltag der Fachkräfte reflektiert werden. Der Fokus liegt ebenfalls auf der Verstärkung von positivem Verhalten. Die statistische Auswertung ergab einen signifikanten Interaktionseffekt, der auf eine bedeutend größere Verbesserung im feinfühligem Verhalten in der Interventionsgruppe hindeutet.

Häufig wird Feedback im pädagogischen Kontext im Rahmen von umfangreichen Coaching- und Qualifizierungsmaßnahmen eingesetzt. Mit „MyTeachingPartner“, dem „Heidelberger Interaktionstraining“ und dem „Hanen-Program“ wurden in Kapitel 2.2.2 bereits wirksame Ansätze vorgestellt, die video-basiertes Feedback verwenden.

Ein dem Feedback verwandtes Konzept ist die Methode der „videostimulierten Reflexion“ (VSR), bei der das eigene pädagogische Handeln kriteriengeleitet von der jeweiligen Person selbst reflektiert wird (Tures, 2014). Im Rahmen von VSR-Interventionen werden die Teilnehmer meist zunächst dazu aufgefordert, das zuvor auf der Videoaufnahme gesehene Geschehen frei zu kommentieren. Im Anschluss daran werden ihnen Reflexionsfragen gestellt, um zu einer eher kriteriengeleiteten Analyse überzugehen (Cutrim Schmid, 2011). Video-basierte Selbstreflexion wird auch im „CAMP Quality Model“ umgesetzt. Dabei geht es um ein umfassendes Fortbildungsformat, welches sich auf die CLASS-Dimensionen bezieht und neben Workshops und Peer-Coaching die Teilnehmer mithilfe verschiedener schriftlich übermittelter Fragen dazu auffordert, das eigenen Handeln zu reflektieren (Zan & Donegan-Ritter, 2014). Bei der VSR-Methode handelt es sich nicht um Feedback im eigentlichen Sinne, da die erreichte Leistung nicht von einer externen Person zurückgemeldet wird. Dennoch ist das Konzept für Feedback-Methoden relevant, da Reflexionsfragen auch im Rahmen von Feedback im Sinne des „scaffolding“ eine aktive Auseinandersetzung mit dem eigenen Verhalten anregen können.

Insgesamt gibt es bisher wenige Studien, die sich konkret mit Feedback für Frühpädagogen beschäftigt und dessen Effekte systematisch untersucht haben (Fukkink & Tavecchio, 2010). Viele Studien haben nur sehr geringe Stichprobengrößen (z.B. Bishop, Snyder &

Crow, 2015; Dennis & Horn, 2013; Englund, 2010), weshalb ihre Ergebnisse nicht empirisch belastbar sind. Die vorhandenen experimentellen Studien zu den Effekten von Feedback deuten auf eine positive Wirkung von Feedback-Interventionen auf das Interaktionsverhalten von Fachkräften hin (z.B. Pianta, Mashburn et al., 2008; Fukkink & Tavecchio, 2010; Hamre et al., 2012; Kaufman, Coddington, Markus, Tryon & Kyse, 2013; Helmerhorst et al., 2017).

### 3.3 Implikationen für die Gestaltung von Feedback

Zusammenfassende Meta-Analysen zur Wirkung von Feedback (z.B. Bangert-Drowns et al., 1991; Hattie & Timperley, 2007; Kluger & DeNisi, 1996) konnten zeigen, dass die Studien eine sehr große Varianz in ihren Effektstärken aufweisen und ein nicht zu vernachlässigender Teil der Untersuchungen sogar negative Effekte von Feedback feststellte. Welche Aspekte von Feedback sich als wirksam oder weniger wirksam erwiesen haben, soll an dieser Stelle zusammenfassend dargestellt werden.

Bezüglich des Informationsgehaltes von Feedback lässt sich festhalten, dass einfache Feedbackformen, welche nur Informationen über das erzielte Ergebnis liefern, niedrigere Effekte haben als elaboriertere Feedbackarten (Bangert-Drowns et al., 1991). Zur Bezugsnorm beim Feedback gibt es unterschiedliche Ergebnisse, wobei in der Feedbackliteratur (z.B. Wollenschläger et al., 2011) insgesamt positive Effekte von sachlich oder individuell begründetem Feedback gegenüber einer sozialen Bezugsnorm überwiegen. Basiert ein Feedback ausschließlich auf Lob hat dieses geringere Effekte (Hattie & Timperley, 2007; Kluger & DeNisi, 1996). Gleiches gilt für extrinsische Belohnungen (Hattie & Timperley, 2007). Feedbackformen, welche konkret bestimmte Lernziele einbeziehen, haben höhere Lerneffekte als andere (Hattie, 1999). Was das Timing von Feedback angeht, gibt es ebenfalls inkonsistente Befunde. Im Großen und Ganzen scheint ein sofortiges Feedback bei Lernenden mit geringerer Kompetenz höhere Effekte zu haben, vor allem wenn die Aufgabe als schwierig empfunden wird. Im Gegensatz dazu ist ein verzögertes Feedback bei kompetenteren Lernern und Aufgaben mit hohem Transfer wirksamer (Hattie & Wollenschläger, 2014). Mit Blick auf den verwendeten Instruktionsansatz scheinen Feedbackinterventionen im Rahmen von programmierten Unterrichtseinheiten eine geringe Wirkung auf das Lernen zu haben (Bangert-Drowns et al., 1991). Dies wird in der Literatur durch das Konzept der „presearch availability“ erklärt, welches aussagt, dass Lernprozesse verhindert werden, wenn Lernende über das Feedback zu früh die korrekte Lösung erfahren (Kulhavy, 1977). Auch die Präsentation der Feedbackinhalte spielt eine wichtige Rolle. So deuten die empirischen Ergebnisse aus der Meta-Analyse von Kluger und DeNisi (1996) zum Beispiel darauf hin, dass eine mündliche Übermittlung zu geringeren Feedbackeffekten führt. Dies



wird von den Autoren unter anderem darauf zurückgeführt, dass bei einem mündlichen Feedback Eigenschaften des Feedbackgebers und deren Repräsentation beim Feedbackempfänger stärker zum Tragen kommen. Auf Seiten des Feedbackempfängers konnte vor allem für das Attributionsverhalten und die Selbstwirksamkeit (z.B. Lyden et al., 2002), sowie die Zielorientierung (z.B. Dahling & Ruppel, 2016) ein Einfluss auf die Feedbackverarbeitung festgestellt werden. Zur Frage, ob computerunterstützte Feedbackformen besser geeignet sind als andere, gibt es keinen eindeutigen Befund. Bei Audio- und Video-Feedback hingegen wurden signifikant höhere Effekte gefunden (Hattie, 1999). Die positive Wirkung von videogestützten Methoden findet sich auch in Studien bei pädagogischen Fachkräften wieder (siehe Kapitel 3.2.6).

Die Ergebnisse der empirischen Studien wurden in verschiedenen Modellen zur Wirkung von Feedback integriert. Das Feedback-Modell von Butler und Winne (1995) beschäftigt sich mit dem selbstregulierten Lernen im Kontext von Feedbacksituationen. Sie gehen davon aus, dass elaborierte Feedbackformen Lernenden dabei helfen können, das eigene Lernen zu steuern. Die „Feedback Intervention Theory“ von Kluger und DeNisi (1996) postuliert, dass die Ziele beim Lernen auf drei hierarchisch angeordneten Ebenen angesiedelt sind: „meta-task processes“ (selbstbezogene Ziele), „task-motivation processes“ (aufgabenbezogene Motivation) und „task-learning processes“ (Steuerung der einzelnen Schritte bei der Aufgabenbearbeitung). Zentrale Annahme ihrer Theorie ist, dass Feedback, welches das Selbst fokussiert, geringere Effekte hat als Feedback, welches die Aufmerksamkeit auf die Motivations- oder Lern-Ebene richtet. Dies kann durch lösungs- oder korrekturrelevante Hinweise erreicht werden, was ebenfalls die Bedeutung von elaboriertem Feedback hervorhebt. Im Mittelpunkt des Modells von Bangert-Drowns et al. (1991) steht das sogenannte „mindful processing“, welches eine zielgerichtete, aktive und intensive Verarbeitung von Informationen meint. Sie gehen davon aus, dass Feedback lernförderlich sein kann, wenn der Lernende zum „mindful processing“ angeregt wird. Auch im Modell von Narciss (2006) spielt die aktive Feedbackverarbeitung durch den Feedbackempfänger eine wichtige Rolle. Ihr „Interactive-Two-Feedback-Loops-Model“ basiert auf der These, dass Feedback nicht ausschließlich eine Vermittlung von evaluativen, ergebnisorientierten Informationen darstellen sollte, sondern den Lernenden beim konstruktiven Prozess des Wissenserwerbs unterstützen muss. Eingebettet in das „visible learning and teaching model“ (Hattie & Timperley, 2007) haben Hattie und Wollenschläger (2014) eine Konzeptualisierung der Wirkungsweise von Feedback vorgenommen. Um die Diskrepanz zwischen einem Ist-Zustand und angestrebten Zielen verringern zu können, muss ein Feedback nach ihrer Auffassung drei zentrale Fragen beantworten: „Where am I going?“ (Ziele und Erfolgskriterien),

„How am I going?“ (aktueller Stand der Leistung) und „Where to next?“ (Wie kann Leistung verbessert werden?).

Die beschriebenen empirischen Befunde sowie die daraus entstandenen Feedback-Modelle liefern somit wichtige Hinweise für die Gestaltung von Feedback. Tabelle 3 gibt einen Überblick über die wichtigsten Aspekte.

Tabelle 3: Implikationen für die Gestaltung von Feedback

Aspekt	Kriterien für die Gestaltung
Feedbackinhalt und -umfang	<ul style="list-style-type: none"> <li>• Feedback kann dazu dienen, die Diskrepanz zwischen einem Ist- und einem Soll-Zustand zu reduzieren. Dazu müssen die Ziele der Aufgabe bzw. die Erfolgskriterien für das erwünschte Verhalten klar kommuniziert und der Lernende über die erbrachte Leistung informiert werden.</li> <li>• Das Feedback zur erbrachten Leistung der Person sollte sich auf ein sachliches (z.B. Kompetenzprofil) oder individuelles Kriterium (z.B. Lernfortschritt) beziehen und nicht auf einen sozialen Vergleich abzielen.</li> <li>• Über ergebnisorientierte Informationen hinaus sollte ein Feedback auch Informationen zur Aufgaben und deren Bearbeitung liefern (elaboriertes Feedback). Wichtig ist es, der lernenden Person Hinweise zu geben, wie das eigene Verhalten bzw. die Leistung verbessert werden kann.</li> <li>• Der Feedbackinhalt sollte nicht die Person des Lernenden fokussieren (z.B. durch allgemeines Lob), sondern das konkrete Verhalten oder die Leistung in Bezug auf eine Aufgabe.</li> <li>• Die Informationsmenge des Feedbacks darf nur so groß sein, dass es den Lernenden nicht überfordert.</li> <li>• Ergänzend zum Feedback können Zielsetzungsmaßnahmen eingesetzt werden.</li> </ul>
Feedback-Präsentation	<ul style="list-style-type: none"> <li>• Das Feedback sollte spezifisch und klar verständlich formuliert sein.</li> <li>• Die Feedbacknachricht sollte so objektiv und neutral wie möglich übermittelt werden, um personale Einflüsse des Feedbackgebers zu minimieren (z.B. durch schriftliches Feedback).</li> <li>• Sofern diese vorhanden und passend sind, können Audio- und Videoelemente unterstützend verwendet werden.</li> <li>• Um „presearch availability“ zu vermeiden, sollte das Feedback erst dann präsentiert werden, wenn der Lernende die Aufgabe bearbeitet hat.</li> <li>• Das Feedback sollte die aktive Konstruktion von Wissen anregen.</li> </ul>
Feedback-Timing	<ul style="list-style-type: none"> <li>• Bei weniger kompetenten Lernern bietet sich ein unmittelbares Feedback an, vor allem wenn die Aufgabe als schwierig empfunden wird.</li> <li>• Bei kompetenten Lernern ist verzögertes Feedback effektiver.</li> <li>• Um Lerntransfer zu unterstützen, ist verzögertes Feedback geeigneter.</li> <li>• Sofortiges Feedback ist auf dem „task level“ (Information über erbrachte Leistung) effektiver, verzögertes Feedback hingegen ist auf dem „process level“ wirksamer.</li> </ul>

## 4 Zusammenfassung, Forschungslücke und Ziele

In Kapitel 2.1.1 wurde ausführlich beschrieben, welche Bedeutung die täglichen Interaktionen mit pädagogischen Fachkräften für die sozial-emotionale, kognitive und sprachliche Entwicklung der Kinder haben. In zahlreichen empirischen Studien (z.B. Mashburn et al., 2008; Siraj-Blatchford et al., 2002; Sylva et al., 2004) wurde der signifikante Zusammenhang zwischen der Qualität der Interaktionen von Fachkräften und Kindern im Kindergarten und der kindlichen Entwicklung bestätigt. Die entwicklungsrelevanten Aspekte von Interaktionen wurden in vielen Beobachtungsverfahren operationalisiert – im Rahmen von Prozessqualität (z.B. ECERS-R), über die Beziehungs- oder Bindungsqualität (z.B. Fremde Situation, AQS, STRS), über das Verhalten der Fachkräfte (z.B. CIS, CLASS, GiNA), oder mit dem Fokus auf bestimmte Aspekte der Entwicklung (z.B. Sprache: DO-RESI, TILRS). Sowohl deutsche als auch internationale Wissenschaftler berichten von einer mittleren bis hohen Qualität im Bereich der emotionalen Unterstützung sowie in der Organisation des Gruppenalltags, sehen jedoch angesichts niedriger Qualitätswerte Entwicklungsbedarf hinsichtlich lernunterstützender Interaktionen von Fachkräften (z.B. Curby et al., 2009; Helmerhorst et al., 2017; Wirts et al., 2017). Die Ergebnisse von Meta-Analysen deuten darauf hin, dass Fort- und Weiterbildungen positive Effekte auf das Interaktionsverhalten von Fachkräften haben können (z.B. Egert, 2015; Werner et al., 2016). Individuelle und praxisnahe Formate, wie zum Beispiel Feedback, konnten dabei als besonders wirksame Methoden identifiziert werden (z.B. Egert, 2015).

Gemäß der allgemeinen Definition von Müller und Ditton (2014, S. 15) wird Feedback in der vorliegenden Arbeit als *„die bewusste Rückmeldung von Informationen an eine Person zu ihrem vorherigen Verhalten“* verstanden. In der heutigen Feedback-Forschung herrscht Konsens darüber, dass in Anlehnung an die systemtheoretische und kybernetische Sichtweise der Nutzen von Feedback darin liegen kann, die Diskrepanz zwischen einem Ist- und einem Soll-Zustand zu minimieren. Im Großen und Ganzen kann zwischen einfachen und elaborierten Feedbackformen unterschieden werden. Im Laufe der Feedbackforschung wurden vor allem drei Wirkungsweisen von Feedback in den Vordergrund gerückt, für die es wissenschaftliche Belege gibt (vgl. Kapitel 3.2.2). Es ist demnach davon auszugehen, dass eine positive Wirkung von Feedback auf der Verstärkung von richtigen Antworten, der Korrektur falscher Lösungen sowie motivationalen Prozessen basiert (Narciss, 2006). Die Feedbackforschung in Lehr-Lernkontexten zeigt, dass Feedback ein effektives Mittel sein kann, um Lernen positiv zu beeinflussen (z.B. Bangert-Drowns et al., 1991; Hattie & Timperley, 2007; Kluger & DeNisi, 1996). Die Ergebnisse der Meta-Analysen machen jedoch auch deutlich, dass die Wirkung von Feedback sehr stark zwischen den Studien variiert, was auf

unterschiedliche Bedingungs- und Gestaltungsfaktoren zurückgeführt werden kann (siehe Kapitel 3.2.4). Im Laufe der Zeit haben verschiedene Autoren versucht, die vielfältigen Befunde zur Wirkung von Feedback in theoretische Modelle zu integrieren, welche wiederum Hinweise für eine effektive Gestaltung von Feedback liefern (siehe Kapitel 3.2.5). Aus den Ergebnissen der dargestellten Studien sowie den Modellen zur Wirkung von Feedback können Implikationen für die Gestaltung von Feedback abgeleitet werden, welche in Kapitel 3.3 zusammengefasst sind.

In Deutschland wird Feedback für frühpädagogische Fachkräfte mittlerweile häufig im Rahmen von Coaching-Programmen (z.B. VIG oder Marte Meo) oder anderen Weiterbildungsformaten (z.B. HIT) eingesetzt. Aber auch Forschungsprojekte können zur Qualitätssicherung und -entwicklung beitragen, indem die gewonnenen Beobachtungsergebnisse an die Praxis zurückgemeldet werden (Hippel & Grimm, 2010). Das Projekt „BiSS-Evaluation“ am Staatsinstitut für Frühpädagogik hat sich genau dies zum Ziel gesetzt. Ein empiriebasierendes, standardisiertes und evaluiertes Konzept für die Rückmeldung von Beobachtungsergebnissen zum sprachlichen Interaktionsverhalten von frühpädagogischen Fachkräften, gibt es bislang für den deutschsprachigen Raum jedoch nicht. Auf der Basis der empirischen Hinweise für die Gestaltung von effektivem Feedback wurde deshalb von der Autorin im Rahmen des genannten Projekts ein Konzept entwickelt, mithilfe dessen frühpädagogischen Fachkräften ressourcenschonend ein Feedback zu ihrem sprachlichen Interaktionsverhalten gegeben werden kann.

Wie beschrieben, haben Weiterbildungsprogramme, die Feedback verwenden, positive Effekte auf das Interaktionsverhalten von frühpädagogischen Fachkräften. Insgesamt gibt es bisher jedoch kaum Studien, die konkret die Effekte von Feedback für Frühpädagogen systematisch untersucht haben und nicht nur als Teil eines umfangreicheren Programms. Die vorhandenen Studien weisen zudem häufig methodische Mängel auf (Fukkink & Tavecchio, 2010). Ziel dieser Studie war daher die empirische Überprüfung der Effekte von Feedback, welches auf der Grundlage des entwickelten Konzepts übermittelt wird.

## 5 Konzept für ein online-basiertes Videofeedback

Bevor auf die Evaluationsstudie eingegangen wird, sollen an dieser Stelle die Grundzüge des in dieser Studie umgesetzten Feedback-Konzepts beschrieben werden. Eine detaillierte Anleitung für die Durchführung eines solchen Feedbackgesprächs und dazu passendes Arbeitsmaterial findet sich bei Kappauf (2018). Für die Entwicklung des Feedback-Konzepts wurde eine umfangreiche Literaturrecherche durchgeführt, deren Ergebnisse in Kapitel 2 und Kapitel 3 dargestellt wurden. Zusätzlich nahm die Autorin im Februar 2016 an einer Schulung zum Thema „Feedback Strategies“ teil. Bei dieser in den USA angebotenen Schulung ging es konkret darum, wie Fachkräften im Kindergarten sinnvoll Beobachtungsergebnisse zurückgemeldet werden können. Das Vorgehen beim Feedback wird jeweils mit den entsprechenden Annahmen bzw. Befunden aus der Literatur begründet. Die Umsetzung des Feedback-Konzepts wurde an einer Stichprobe erprobt – Umsetzung und Ergebnisse werden im letzten Abschnitt dieses Kapitels zusammengefasst.

### 5.1 Eckpfeiler des Feedback-Konzepts

Zunächst werden einige Eckpunkte des Feedback-Konzepts thematisiert. Diese beziehen sich auf zentrale Aspekte der Feedbackgestaltung (siehe Tabelle 3, S. 69)

#### Feedback-Präsentation

Nach Narciss (2014) sowie auch Bangert-Drowns et al. (1991) sollte Feedback die bewusste und aktive Konstruktion von Wissen anregen und nicht nur ergebnisorientierte Informationen anbieten. Im Sinne dieses ko-konstruktivistischen Ansatzes soll im Rahmen des Feedbacks das Interaktionsverhalten der Fachkraft gemeinsam mit dieser reflektiert werden. Ziel ist es zudem, jede Fachkraft „abzuholen“, also ihre individuellen Voraussetzungen (z.B. vorhandenes Wissen) zu berücksichtigen. Aus diesen Gründen wurde ein mündliches Feedbackgespräch einem schriftlichen Feedback vorgezogen. Ein persönliches Feedbackgespräch beansprucht jedoch hohe personale und finanzielle Ressourcen. Das Feedback wurde daher online mithilfe einer Videokonferenzsoftware durchgeführt.

Das Feedback sollte zudem so gestaltet sein, dass eine Verknüpfung der Feedbackinhalte mit dem eigenem Verhalten möglich ist (Antons, 2000). Videogestützte Formate haben sich als besonders effektiv erwiesen (z.B. Hattie, 1999). Das Feedback wurde deshalb videobasiert konzipiert und bezieht sich auf Videoaufnahmen, welche im Alltag der Fachkräfte angefertigt wurden. Auf diesem Weg haben die Fachkräfte die Möglichkeit, das eigene Ver-

halten noch mal in Ruhe „von außen“ zu betrachten und müssen sich nicht auf ihre Erinnerung berufen.

Für die technische Umsetzung des online-basierten Videofeedbacks wurde eine umfangreiche Recherche und Testung von potenziellen Softwarelösungen vorgenommen. Da über die Software Videomaterial abgespielt werden sollte, muss diese einen sehr hohen Datenschutz garantieren. Dies konnte bei den meisten Lösungen nicht garantiert werden, da viele aus dem amerikanischen Raum stammen und ihren Serverstandort nicht in Deutschland haben, weshalb nicht sichergestellt werden kann, dass sie nach dem Bundesdatenschutzgesetz (BDSG) arbeiten (z.B. Skype, WebEx, IBM Sametime, Teamviewer, Citrix GoToMeeting). Außerdem mussten spezielle technische Möglichkeiten umgesetzt werden können. Grundlegende Voraussetzung war, dass die Videoaufnahmen während des Gesprächs abgespielt werden können. Einige Anbieter bieten diese Funktion gar nicht an (z.B. Patientus), bei anderen Softwarelösungen kam es in der Testphase zu technischen Problemen (z.B. kein Ton). Wie in Kapitel 2.1.1 angedeutet, spielen Mimik und Gestik eine bedeutende Rolle in der Kommunikation mit Menschen. Für das Feedback wurde es daher als hilfreich erachtet, dass man den Gesprächspartner während des Online-Gesprächs sehen kann. Es wurde dementsprechend vorausgesetzt, dass während des Gesprächs eine Webcam-Übertragung möglich ist. Ein deutsches Unternehmen, welches die strengen Datenschutzrichtlinien sowie die erforderlichen technischen Voraussetzungen erfüllt, ist die vitero GmbH. Vitero ist eine Videokonferenzsoftware, die auf E-Learning-Angebote spezialisiert ist. Die Daten werden nur verschlüsselt übertragen. Die Videoaufnahmen liegen nur bis zur Löschung durch den Benutzer in der Cloud und werden darüber hinaus nicht auf einem der Server gespeichert. Die vitero-Mitarbeiter haben keinen Zugang zu den Videoaufnahmen. Die pädagogischen Fachkräfte können die Videosequenzen während des Gesprächs sehen, haben aber keinen Zugriff darauf und können diese auch nicht auf ihrem Computer speichern. Obwohl vitero kostenpflichtig ist, wurde die Software für die Feedbackgespräche verwendet, da keine vergleichbar gute kostenfreie Lösung zur Verfügung stand.

### Feedbackinhalt und -umfang

Da in einem Feedbackgespräch immer nur eine begrenzte Menge an Inhalt effektiv thematisiert werden kann und sollte (Shute, 2008), wird der Inhalt und somit auch der Umfang des Feedbacks auf zwei Ebenen eingegrenzt: Situation und inhaltlicher Fokus.

Der Alltag im Kindergarten ist sehr vielseitig, weshalb das Interaktionsverhalten in sehr unterschiedlichen Alltagssituationen beobachtet werden kann (z.B. Morgenkreis, Freispiel oder Essen). Innerhalb eines Feedbackgesprächs können jedoch nicht alle Situationen sinnvoll abgedeckt werden. Aus diesem Grund bezieht sich das Feedback-Konzept beispielhaft auf die Situation der Bilderbuchbetrachtung. Eine Anpassung auf andere Situatio-

nen ist allerdings gut möglich. Die Bilderbuchbetrachtung wurde vor allem deshalb ausgewählt, weil diese Aktivität auch in der Kleingruppe durchgeführt werden kann und daher nicht zwangsweise die Einwilligung aller Eltern in der Kindergartengruppe für die Videoaufnahmen erforderlich ist.

Auf der inhaltlichen Ebene können sehr unterschiedliche Aspekte des Interaktionsverhaltens thematisiert werden (vgl. Kapitel 2.1.2). Da das Feedback-Konzept ursprünglich für Rückmeldungen im Rahmen des Projekts „BiSS-Evaluation“ entwickelt wurde, bezieht sich das Feedback-Konzept auf das sprachliche Interaktionsverhalten der Fachkräfte.

Ein Feedback kann dazu dienen, die Diskrepanz zwischen einem Ist-Zustand und einem Soll-Zustand zu minimieren (vgl. Kapitel 3.1.1 und Kapitel 3.2.5). Innerhalb eines Feedbacks kann der Ist-Zustand mit einem sachlich begründeten Lernkriterium, mit der vorangegangenen Leistung oder mit der Leistung anderer verglichen werden. Der Ist-Zustand stellt in diesem Fall das sprachanregende Interaktionsverhalten der Fachkraft dar. Die Ergebnisse verschiedener Studien (vgl. Kapitel 3.2.4) weisen darauf hin, dass es motivationsfördernder ist, wenn sich das Feedback auf ein sachliches oder individuelles Kriterium bezieht als auf ein soziales Kriterium. Da für eine individuelle Bezugsnormorientierung immer mehrere Beobachtungszeitpunkte notwendig wären, nimmt das Feedback hier ausschließlich Bezug auf ein sachliches Kriterium. Bei der Einschätzung des sprachlichen Verhaltens (Ist-Stand) sollte sichergestellt sein, dass es sich um für die kindliche Entwicklung relevante Kriterien handelt und andere Feedbackgeber zu einer ähnlichen Bewertung kommen würden. Da das „Classroom Assessment Scoring System“ (Pianta, La Paro et al., 2008) weltweit in vielen Studien erprobt wurde, in denen auch signifikante Zusammenhänge zur kindlichen Entwicklung aufgezeigt wurden (vgl. Kapitel 2.1.2 und Kapitel 2.1.3), wurden für die vorliegende Arbeit vor allem die Indikatoren der Dimension „Language Modeling“ aus der CLASS für den Ist-Soll-Vergleich herangezogen. Das Feedback-Konzept kann jedoch auch für andere Aspekte des Interaktionsverhaltens oder andere Beobachtungsverfahren eingesetzt werden.

### Feedback-Timing

Da das Feedback auf Videoaufnahmen beruht, die zunächst ausgewertet und aufbereitet werden müssen, ist eine gewisse Vorbereitungszeit nötig. Ein unmittelbares Feedback ist aus diesem Grund nicht möglich. Das Konzept ist deshalb so angelegt, dass das Feedbackgespräch erst verzögert mit einem gewissen Zeitabstand zur Videoaufnahme durchgeführt wird.



## 5.2 Vorgehen beim Feedbackgespräch

In Kapitel 3.2.5 wurden verschiedene Feedback-Modelle beschrieben. Diese stellen aus unterschiedlichen Perspektiven die Wirkung von Feedback in Abhängigkeit von verschiedenen Einflussfaktoren dar und liefern so einen guten Rahmen für die Gestaltung von Feedback. Das Feedback-Modell von Hattie und Timperley (2007) bzw. Hattie und Wollenschläger (2014) stellt den komplexen Prozess der Feedbackwirkung multidimensional dar. Dabei ist es so allgemein und dennoch praxisnah formuliert, dass es sich gut auf die pädagogische Arbeit von Fachkräften im Kindergarten und die Gestaltung von Feedback in diesem Kontext beziehen lässt. Das vorliegende Feedback-Konzept lehnt sich daher in seinen Grundzügen an das Modell von Hattie und Timperley (2007) an, wobei die empirischen Befunde zu den Wirkbedingungen von Feedback und die darauf aufbauenden Annahmen der übrigen Feedback-Modelle ebenfalls Berücksichtigung finden. Es besteht vor allem aus einem detaillierten Leitfaden und einer Arbeitsvorlage. Der Leitfaden enthält zahlreiche beispielhafte Formulierungen, um die konkrete Umsetzung zu erleichtern und gleichzeitig zu standardisieren. Es wird beispielhaft gezeigt, wie der einzelne Schritt im Feedbackgespräch eingeführt werden kann. Zusätzlich gibt es Beispielformulierungen für die Darstellung von Kriterien guter sprachlicher Anregung, die Rückmeldung der Beobachtungsergebnisse und Reflexionsfragen an die Fachkraft. Basierend auf dem Leitfaden wurde zudem eine Arbeitsvorlage entworfen, mithilfe derer das Feedbackgespräch Schritt für Schritt vorbereitet und durchgeführt werden kann. In dieser gibt es zu jedem Baustein Textfelder, in denen eigene Formulierungen und die jeweiligen individuellen Beispiele für das Feedbackgespräch festgehalten werden können. Im Folgenden werden die einzelnen Schritte, wie sie in der vorliegenden Arbeit umgesetzt wurden, zusammenfassend vorgestellt. Das Vorgehen wird ausführlich in einer Broschüre (Kappauf, 2018) beschrieben, die darüber hinaus Kopiervorlagen zur individuellen Anwendung beinhaltet.

### 5.2.1 Vorbereitung des Feedbackgesprächs

Um eine möglichst hohe Qualität des Feedbackgesprächs zu gewährleisten, bedarf es einer guten Vorbereitung. Unabdingbare Voraussetzung ist, dass eine intensive Auseinandersetzung mit dem Material erfolgt. Mit der Arbeitsvorlage kann schließlich konkret das jeweilige Feedbackgespräch vorbereitet werden. Die einzelnen Schritte werden im Folgenden skizziert.

#### Einschätzung des sprachlichen Interaktionsverhaltens

Wenn die Videoaufnahme vorliegt, muss diese zunächst hinsichtlich der sprachlichen Unterstützung analysiert werden. Wie bereits beschrieben, wurde für diese Studie dafür die

Dimension „Language Modeling“ aus der CLASS verwendet. Das sprachanregende Verhalten der Fachkraft wurde anhand der folgenden fünf Indikatoren beschrieben: „Frequent conversation“, „Open-ended questions“, „Repetition and extension“, „Self- and parallel talk“ sowie „Advanced Language“. Eine ausführliche Beschreibung der Indikatoren findet sich in Kapitel 6.3.1. Für das Feedbackgespräch ist es weniger wichtig, einen Wert für die Qualität der sprachlichen Anregung zu vergeben, als vielmehr ausführliche Notizen zu den einzelnen Indikatoren anzufertigen. Nur mithilfe konkreter Beispiele kann der Fachkraft verdeutlicht werden, was schon sehr gut gelingt und an welchen Stellen es noch Optimierungspotenzial gibt. In der Arbeitsvorlage werden deshalb Beispiele für die einzelnen Indikatoren (z.B. längere Gespräche oder offene bzw. geschlossene Fragen) mit der entsprechenden Zeitangabe im Video notiert.

### Auswahl der Indikatoren

Auf Basis dieser Notizen werden anschließend drei Indikatoren ausgewählt, die im Feedbackgespräch thematisiert werden sollen. Bei der Auswahl der Indikatoren sollten folgende Kriterien berücksichtigt werden:

1. Um die Fachkraft in ihrem positiven Verhalten zu bestärken, wird als Erstes der Indikator gewählt, den die Fachkraft am besten umgesetzt hat. Wenn nicht genug positive Beispiele zur sprachlichen Anregung vorhanden sind, werden andere Aspekte (z.B. positive Atmosphäre, Gestaltung der Situation o.ä.) an dieser Stelle genannt.
2. Als nächstes sollte ein Indikator besprochen werden, der in manchen Situationen bereits gut umgesetzt wird, aber noch häufiger/konsequenter gezeigt werden könnte.
3. Auch der dritte Indikator sollte in manchen Situationen bereits gut umgesetzt werden, aber noch Optimierungspotenzial haben. Wenn es keinen zweiten Indikator mit positiven Beispielen gibt, wird ein Indikator reflektiert, der noch gar nicht umgesetzt wird.

Da Dialoge mit Kindern eine wesentliche Voraussetzung für deren sprachliche Entwicklung sind (vgl. Kapitel 2.1.1 und Kapitel 2.1.3), sollte dieser Aspekt auf jeden Fall thematisiert werden. Das handlungsbegleitende Sprechen hat bei der Bilderbuchbetrachtung nur eine zweitrangige Bedeutung, da es in der Regel sehr wenige Handlungen gibt, die sprachlich begleitet werden können. Es kann deshalb für das Feedback in Bezug auf eine Bilderbuchbetrachtung gegenüber den anderen Indikatoren vernachlässigt werden. Das besprochene Optimierungspotenzial der Fachkraft sollte innerhalb ihrer „Zone der nächsten Entwicklung“ liegen. Wenn eine Fachkraft beispielsweise nur vorliest und gar nicht mit den Kindern spricht, ist es nicht hilfreich direkt bei längeren Dialogen anzufangen, sondern z.B. erst mal

mit dem Stellen von Fragen. Wenn sich in der Aufnahme andere Aspekte des Interaktionsverhaltens zeigen, die unbedingt eine Rückmeldung benötigen (z.B. aggressives Verhalten o.ä.), sollten diese zusätzlich in das Feedback einbezogen werden.

### Auswahl der Videosequenzen

Für jeden Indikator müssen schließlich Videosequenzen ausgewählt werden, die passende Verhaltensbeispiele enthalten. Jede Sequenz sollte im Idealfall nicht länger als zwei Minuten dauern. In der Regel reichen bereits 30 Sekunden bis eine Minute aus, je nach Situation können aber auch längere Sequenzen Sinn machen. Dabei sollte darauf geachtet werden, dass sie nicht nur das entsprechende Beispiel zeigen, was besprochen werden soll, sondern auch Handlung davor und danach. Nur so kann sichergestellt werden, dass die Fachkraft weder durch zu viele Beispiele überfordert noch durch zu wenig Reflexion unterfordert wird. Insgesamt sollten die Videosequenzen idealerweise positive Beispiele für Indikator 1 sowie positive Beispiele und Beispiele mit Optimierungspotenzial für Indikator 2 und 3 enthalten. Wenn es möglich ist, sollte zusätzlich eine längere Sequenz ausgewählt werden, welche das Verhalten der Fachkraft gut widerspiegelt. Wichtig ist, dass diese Gesamtsequenz auf jeden Fall auch positive Verhaltensbeispiele zeigt. Im optimalen Fall enthält die Gesamtsequenz zumindest einen Teil der Einzelsequenzen. Manchmal kann es auch Sinn machen, eine Gesamtsequenz aus mehreren Einzelsequenzen zusammenzuschneiden. Insgesamt sollte sie nicht viel länger als fünf Minuten dauern.

In der Arbeitsvorlage werden die ausgewählten Indikatoren, die dazugehörigen Videosequenzen sowie passende Verhaltensbeispiele der Fachkraft für jeden Indikator notiert.

### Vorbereitung der Sitzung

Sobald das Feedbackgespräch inhaltlich aufbereitet ist, kann die Online-Sitzung vorbereitet werden. Sowohl auf dem Computer des Feedbackgebers als auch auf dem der Fachkraft muss die entsprechende Software installiert sein. Für die hier eingesetzte vitero-Software muss dazu eine „Exe-Datei“ heruntergeladen und ausgeführt werden. Für das Speichern der Datei sind keine Administratorrechte für den Computer notwendig. Dies stellt eine wichtige Voraussetzung dar, da Fachkräfte häufig nur eingeschränkte Nutzungsrechte für die Computer in der Kita haben. Vitero kann schließlich über einen Link geöffnet werden. Der Feedbackgeber meldet sich als Moderator für die Sitzung an. Sobald der Moderator im „virtual team room“ anwesend ist, erhält die Fachkraft Zugang. Die Anmeldung der Fachkraft erfolgt mithilfe eines Sitzungscode und einem persönlichen Passwort. Eine ausführliche Anleitung für vitero wurde sowohl für Fachkräfte als auch für Feedbackgeber formuliert. Alle notwendigen Informationen (Zugangsdaten und Anleitung für den „virtual team room“ von vitero) erhält die Fachkraft vorher per E-Mail. Zusätzlich muss sichergestellt

werden, dass die Fachkraft über die nötige technische Ausstattung (Webcam und Headset) verfügt. Als letzter Schritt vor dem Feedbackgespräch müssen die Videodateien sowie benötigte Unterlagen innerhalb des „Vitero Management Systems“ abgespeichert werden, um während des Gesprächs darauf zugreifen zu können.

## 5.2.2 Durchführung des Feedbackgesprächs

Nach Hattie und Timperley (2007) müssen während eines Feedbacks drei grundlegende Fragen beantwortet werden: „Where am I going?“, „How am I going?“ und „Where to next?“. Die Beantwortung dieser Fragen steht daher im Mittelpunkt des Feedbackgesprächs. Da verschiedene Studien zeigen, dass Zielsetzungsmaßnahmen den Effekt von Feedback steigern (z.B. Kluger & DeNisi, 1996), gibt es zusätzlich einen vierten Baustein, in dem am Ende des Feedbackgesprächs mit der Fachkraft konkrete Ziele schriftlich festgehalten werden. Die einzelnen Bausteine werden in den folgenden Abschnitten vorgestellt.

### Baustein 1: Erfolgskriterien kommunizieren („Where am I going?“)

Bevor mit dem eigentlichen Feedback begonnen werden kann, müssen die Kriterien für eine erfolgreiche Umsetzung des angestrebten Verhaltens (Soll-Zustand) besprochen werden. Zu Beginn des Feedbackgesprächs werden daher die einzelnen Indikatoren für ein erfolgreiches Handeln vorgestellt. Zur Veranschaulichung wurde im Rahmen der vorliegenden Arbeit der Fachkraft innerhalb des „virtual team rooms“ eine Übersicht über die einzelnen Indikatoren gezeigt (vgl. Abbildung 3).

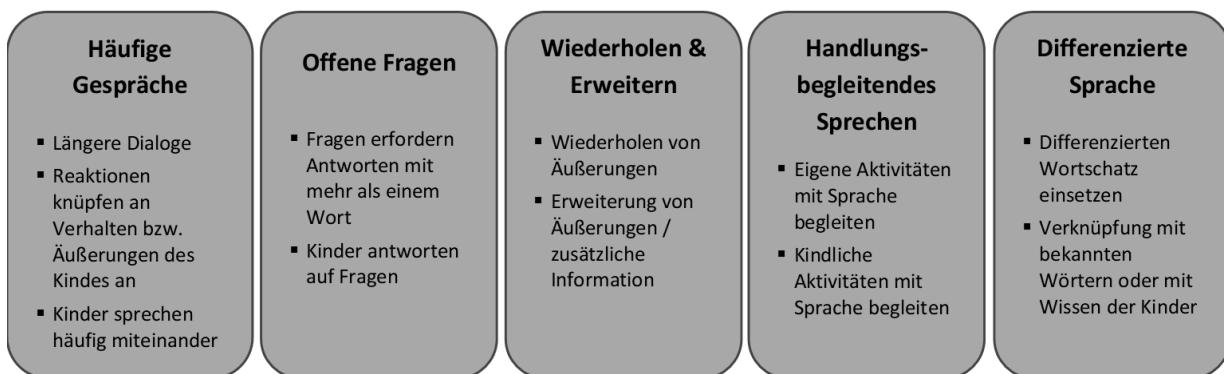


Abbildung 3: Indikatoren der Dimension "Language Modeling" der CLASS für ein sprachanregendes Verhalten

Die Indikatoren wurden nacheinander präsentiert und beschrieben. Um sicherzustellen, dass die Indikatoren möglichst einheitlich und gut verständlich dargestellt werden, gibt es im Arbeitsmaterial eine Beschreibung der Kriterien von guter sprachlicher Anregung.

## Baustein 2: Analyse des „Ist-Stands“ („How am I going?“)

Bevor die einzelnen Schritte dieses Bausteins beschrieben werden, sollen kurz allgemeine Aspekte des Vorgehens thematisiert werden.

Für die Rückmeldung der Beobachtungsergebnisse (Ist-Stand) wurden auf Basis empirischer Ergebnisse aus der Feedback-Forschung (vgl. Kapitel 3.2.4) folgende Richtlinien formuliert:

- Verschiedene Studien weisen auf geringe bzw. negative Effekte von allgemeinem Lob hin (z.B. Kluger & DeNisi, 1996). Dieses sollte deshalb nur geringfügig eingesetzt werden, um die Motivation zu steigern oder das Selbstwertgefühl zu stärken.
- Aufgrund von Hinweisen darauf, dass eine soziale Bezugsnorm motivationshemmend ist (z.B. Krampen, 1987), sollten keine Vergleiche mit Kollegen o.ä. angestellt werden.
- Da eine Bewertung in Form von Werten keine Informationen dazu liefert, welche Aspekte schon gut umgesetzt werden und wo es noch Optimierungspotenzial gibt, wird als Kriterium für den Ist-Stand nicht der Qualitätswert verwendet, sondern das konkrete Verhalten der Fachkraft.
- Dabei darf das Feedback nicht die Person selbst fokussieren, sondern muss sich sachlich auf die gezeigte Performanz beziehen (Hattie & Wollenschläger, 2014).
- Um die Motivation zu steigern und die Entstehung von positiven Selbstwirksamkeitserwartungen zu unterstützen, sollte das Feedback eine internale und veränderbare Attribuierung anregen (Positives Verhalten wird auf eigene Leistung und Anstrengung zurückgeführt, Verhalten kann durch Anstrengung verändert werden).

Nachdem die einzelnen Indikatoren sprachlicher Anregung besprochen wurden, startet die Reflexion der ausgewählten Videosequenzen, deren Ablauf im Folgenden dargestellt wird.

### *1. Allgemeiner Einstieg*

Zunächst wird die ausgewählte Gesamtsequenz gemeinsam mit der Fachkraft angeschaut. Wenn es gar keine passende Gesamtsequenz gibt, wird die Einzelsequenz mit dem Indikator, der am besten umgesetzt wird, gezeigt. Um einen Kontakt zur Fachkraft aufzubauen und einen ersten Eindruck zu bekommen, wird die Fachkraft im Anschluss ganz offen nach ihrer Wahrnehmung der Videosequenz gefragt (z.B. „Wie war es für Sie, diese Sequenz anzuschauen?“). In der Arbeitsvorlage kann die Antwort der Fachkraft notiert werden. Es sollte darauf geachtet werden, dass das Gespräch nicht zu stark auf Nebenaspekte, wie z.B. Aussehen, Stimme etc. oder nur auf negative Aspekte gelenkt wird. Im Arbeitsmaterial werden mögliche Reaktionen für verschiedene Szenarien vorgeschlagen.

### 2. Indikator 1

Das Feedback soll sowohl eine korrigierende als auch eine verstärkende und motivierende Funktion erfüllen. Im Sinne der Verstärkung wird zu Beginn des Feedbacks ein Indikator reflektiert, der während der Bilderbuchbetrachtung gut bzw. am besten umgesetzt wurde. Es kann auch ein anderer Aspekt herangezogen werden (z.B. positive Atmosphäre, Blickkontakt etc.), wenn es keinen gut umgesetzten Indikator für sprachanregendes Verhalten gibt.

Die positiven Beispiele können entweder während des Anschauens der Videoaufnahme oder im Anschluss reflektiert werden. Bei längeren Sequenzen, vielen Beispielen und konkreten Verhaltensbeispielen, bietet es sich an, während der Aufnahme zu reflektieren. Bei kurzen Sequenzen, wenigen Beispielen und eher schwer zu definierenden Beispielen können die Aspekte auch nach der Beispielsequenz besprochen werden.

Wie bereits in Kapitel 5.1 aufgeführt, sollte ein Feedback so gestaltet sein, dass der Feedbackempfänger nicht nur eine Rückmeldung über die eigene Leistung erhält, sondern aktiv einbezogen wird. Auch in der Meta-Analyse von Egert (2015) stellten sich Interventionen, die aktive Lernelemente einsetzten, als effektiver heraus. Das Feedback ist daher so aufgebaut, dass die Videosequenzen gemeinsam mit der Fachkraft reflektiert werden. Wenn während der Aufnahme reflektiert wird, kann man die Fachkraft z.B. bitten, immer „Stopp“ zu sagen, wenn sie/er ein Beispiel für das vorher beschriebene Verhalten entdeckt (z.B. „Bitte sagen Sie immer Stopp oder geben mir ein Zeichen, wenn sie eine Situation erkennen, in der Sie eine offene Frage gestellt haben.“). Alternativ (z.B. wenn die Fachkraft die Beispiele nicht selbst erkennt), drückt man selbst auf „Stopp“ und gibt eine spezifische positive Rückmeldung. Wenn man im Anschluss an die Aufnahme reflektiert, kann man die Fachkraft nach Beispielen für das vorher beschriebene Verhalten in der Sequenz fragen oder selbst eine spezifische positive Rückmeldung geben.

Im Leitfaden sind Beispielformulierungen für mögliche Reflexionsfragen oder positive Rückmeldungen aufgeführt. So könnte die Fachkraft beispielsweise nach der Videosequenz gefragt werden, an welcher Stelle sie die Äußerung eines Kindes wiederholt hat oder ob sie eine Situation benennen kann, in der sie einen längeren Dialog mit einem Kind hatte. Eine positive Rückmeldung in Bezug auf den Aspekt „Anknüpfen“ könnte zum Beispiel wie folgt lauten: „Mit Ihrer Frage danach, welche Tiere Tom bereits gesehen hat, nehmen Sie Bezug auf seine Äußerung, dass er auch schon viele Tiere kennt, und zeigen so Interesse an dem, was er erzählt hat“. In der Arbeitsvorlage werden alle positiven Beispiele der Sequenz notiert, um in der Situation unmittelbar reagieren zu können.

### 3. Indikator 2

Im Anschluss an die Reflexion der gut umgesetzten Beispiele wird ein Indikator besprochen, bei dem die Umsetzung zum Teil schon gelingt, es aber noch Optimierungspotenzial gibt. Dabei wird immer erst mit einem oder mehreren positiven Beispielen für die Umsetzung begonnen. Dadurch hat die Fachkraft eine konkrete Vorstellung vom entsprechenden Indikator und weiß, wie dieser umgesetzt werden kann. Im Anschluss wird besprochen, in welchen Situationen der entsprechende Indikator auch noch eingesetzt werden könnte.

Um sicherzustellen, dass der Fachkraft die Kriterien für den entsprechenden Indikator präsent sind, kann dieser zunächst noch einmal kurz zusammengefasst werden (siehe Baustein 1). Danach wird (werden) die ausgewählte(n) Sequenz(en) mit der Fachkraft angeschaut. Die Beispiele für eine positive Umsetzung und die, in denen das Zielverhalten noch häufiger eingesetzt werden könnte, können entweder in einer Sequenz oder in mehreren Einzelsequenzen hintereinander gezeigt werden.

In Bezug auf die positiven Beispiele für die Umsetzung wird genauso vorgegangen wie beim ersten Indikator. Die Fachkraft solle auf jeden Fall einbezogen werden. Als Ergänzung oder Hilfestellung können positive Beispiele genannt werden. Auch hier werden im Vorhinein alle positiven Beispiele im Rahmen der Vorbereitung des Feedbackgesprächs in der Arbeitsvorlage notiert.

Nach dem positiven Beispiel für die Umsetzung wird die oder (eine) andere Sequenz(en) in Bezug auf ihr Optimierungspotenzial reflektiert. Dabei wird mit der Fachkraft zunächst erarbeitet, an welchen Stellen das sprachunterstützende Verhalten noch umgesetzt werden könnte. Zur Vereinfachung kann man zunächst allgemein besprechen, wie man es schaffen könnte, den entsprechenden Indikator umzusetzen (z.B. häufiger Gespräche führen) und im Anschluss daran die konkrete Umsetzung anhand der Videosequenz erörtern. An dieser Stelle sollte in erster Linie die Fachkraft die dargestellten Situationen reflektieren. Dies kann, wie oben beschrieben, entweder während der Beispielsequenz oder im Anschluss daran erfolgen. Im Leitfaden sind wieder verschiedene Fragen beispielhaft formuliert, mit denen die Fachkraft zur Reflexion ermutigt werden kann. In Bezug auf das Stellen von offenen Fragen könnte man eine Reflexion wie folgt anregen: „Nennen Sie mir ein Beispiel für eine offene und eine geschlossene Frage, die Sie gestellt haben. Vergleichen Sie die Reaktion der Kinder auf die beiden Fragen. Was fällt Ihnen auf?“. Als Ergänzung oder Hilfestellung kann man selbst Hinweise geben oder Beispiele nennen. Dazu sollten, wie bisher, eigene Beispiele als Back-up vorher in der Arbeitsvorlage notiert werden.

Nach Hattie und Wollenschläger (2014) darf ein Feedback nicht nur Optimierungspotenzial benennen, sondern muss auch darüber informieren, wie man sich verbessern kann (Baustein 3: „Where to next?“). Sobald klar ist, an welchen Stellen das entsprechende Verhalten

noch eingesetzt werden könnte, wird daher mit der Fachkraft erarbeitet, wie die Umsetzung auch in diesen Situationen gelingen kann. Die Bausteine 2 und 3 erfolgen demnach in der Umsetzung in einem Schritt. Um das Konzept in seinen Grundrissen beschreiben zu können, wird Baustein 3 später gesondert dargestellt.

### 4. Indikator 3

Die Reflexion der Videosequenzen endet mit einem zweiten Indikator, bei dem es Entwicklungspotenzial gibt. Wenn es sowohl positive Beispiele als auch Situationen mit Optimierungspotenzial im Videomaterial gibt, wird das Vorgehen von Indikator 2 wiederholt. Wenn es keinen zweiten Indikator mit positiven Beispielen für die Umsetzung gibt, wird mit kleinen Anpassungen wie beim Optimierungspotenzial von Indikator 2 vorgegangen. Zunächst wird der Fachkraft noch einmal kurz der Indikator beschrieben, der nun reflektiert werden soll. Während man die Videosequenz(en) anschaut oder im Anschluss daran wird schließlich mit der Fachkraft erarbeitet, wann das entsprechende Verhalten umgesetzt werden könnte. Dabei kann die Fachkraft zum Beispiel danach gefragt werden, an welchen Stellen sie die Möglichkeit gehabt hätte, mit den Kindern ins Gespräch zu kommen, welche Fragen anders formuliert werden könnten oder ob es Gelegenheiten gab, den Wortschatz der Kinder zu erweitern. Falls die Fachkraft die Situationen nicht selbst benennen kann, können Hinweise oder Beispiele gegeben werden, die vorher in der Arbeitsvorlage notiert wurden. Auch hier wird wieder direkt diskutiert, wie eine positive Umsetzung gelingen könnte (siehe Baustein 3).

### Baustein 3: Lösungen erarbeiten („Where to next?“)

Wie bereits erwähnt, ist der dritte Baustein in der Umsetzung in den zweiten Baustein integriert, das heißt: Sobald Situationen oder Beispiele besprochen werden, in denen die Umsetzung noch nicht vollständig gelingt, muss auch gemeinsam mit der Fachkraft erarbeitet werden, wie eine gute Umsetzung gelingen kann. Zum Teil geschieht dies automatisch in einem Schritt, das heißt, wenn die Fachkraft nach Situationen mit Optimierungspotenzial gefragt wird, ist eine mögliche positive Umsetzung manchmal bereits in ihrer Antwort enthalten (Beispiel: Feedbackgeber fragt „An welchen Stellen könnten Sie den Kinder noch mehr offene Fragen stellen?“ und die Fachkraft antwortet: „In dieser Situation könnte ich fragen, wie die Geschichte weiter gehen könnte.“). Wenn dies nicht der Fall ist, kann die Fachkraft durch Reflexionsfragen dazu angeregt werden, über mögliche Lösungen nachzudenken (z.B. „Wie könnten Sie diese Frage offen formulieren?“, „Wie könnte an dieser Stelle eine Antwort von Ihnen lauten, die Bezug auf das Gesagte der Kinder nimmt?“ oder „Was könnten Sie machen, um den Kindern diese neuen Wörter verständlich zu machen?“). Wei-



tere Fragen sind im Leitfaden aufgeführt. Als Back-up sollten mögliche Lösungsvorschläge wieder im Zuge der Vorbereitung in der Arbeitsvorlage notiert werden.

#### Baustein 4: Zielvereinbarung

Da sich in verschiedenen Studien herausgestellt hat, dass Zielsetzungen die Effekte von Feedback erhöhen (Hattie, 1999), wurde zusätzlich zum Modell von Hattie und Timperley (2007) ein vierter Baustein in das Feedback-Konzept integriert. Dieser besteht aus einer Zielvereinbarung.

Basierend auf der Reflexion der Videosequenzen werden am Ende des Feedbackgesprächs konkrete Ziele vereinbart. Dazu wird die Fachkraft gefragt, welche Ziele sie sich auf der Grundlage von dem, was gerade besprochen wurde, setzen will. Für die Formulierung der Ziele muss berücksichtigt werden, wo die Fachkraft gerade steht (Teachstone, 2014). Dabei können drei Ebenen von Zielen unterschieden werden:

- Wissen: Indikatoren sprachlicher Anregung wiederholen
- Reflektieren: Eigenes Verhalten im Alltag reflektieren (Wie viele Gespräche habe ich mit den Kindern? Wie viele offene Fragen stelle ich? etc.)
- Handeln: Erwünschtes Verhalten im Alltag üben (z.B. mehr offene Fragen bei der Bilderbuchbetrachtung stellen etc.)

Oberstes Ziel ist immer die Ebene des Handelns, da so Anwendung und Transfer am besten unterstützt werden. Diese Ebene sollte, wenn möglich, bei der Zielformulierung Berücksichtigung finden. Wenn die Fachkraft jedoch noch nicht soweit ist (z.B. wenn nötiges Wissen fehlt oder die Fachkraft keinen Handlungsbedarf sieht), können die anderen Ebenen vorangestellt werden.

Die vereinbarten Ziele wurden in einer Word-Vorlage schriftlich festgehalten. Anschließend wurde das Dokument unter dem Code der Fachkraft neu abgespeichert und über den „virtual team room“ an die Fachkraft geschickt.

#### Positiver Abschluss

Zum Abschluss des Gesprächs wird noch einmal zusammengefasst, was während des Feedbacks besprochen wurde. Hier sollten unbedingt positive Aspekte berücksichtigt werden.

### **5.2.3 Nachbereitung des Feedbackgesprächs**

Nach dem Feedbackgespräch wurde in der vorliegenden Studie der Fachkraft eine ausführliche Darstellung der Indikatoren für sprachanregendes Verhalten über den „virtual

team room“ oder per E-Mail geschickt. Darüber hinaus wurden folgende Punkte protokolliert:

- Indikatoren, die thematisiert wurden
- Vereinbarte Ziele
- Persönliche Einschätzung des Feedbackgesprächs
- Technische Probleme

## 5.3 Das Feedback-Konzept auf einen Blick

Die folgende Abbildung gibt einen Überblick über die einzelnen Bausteine und Arbeitsschritte des hier eingesetzten Feedbacks, welche in den vergangenen Abschnitten beschrieben wurden.

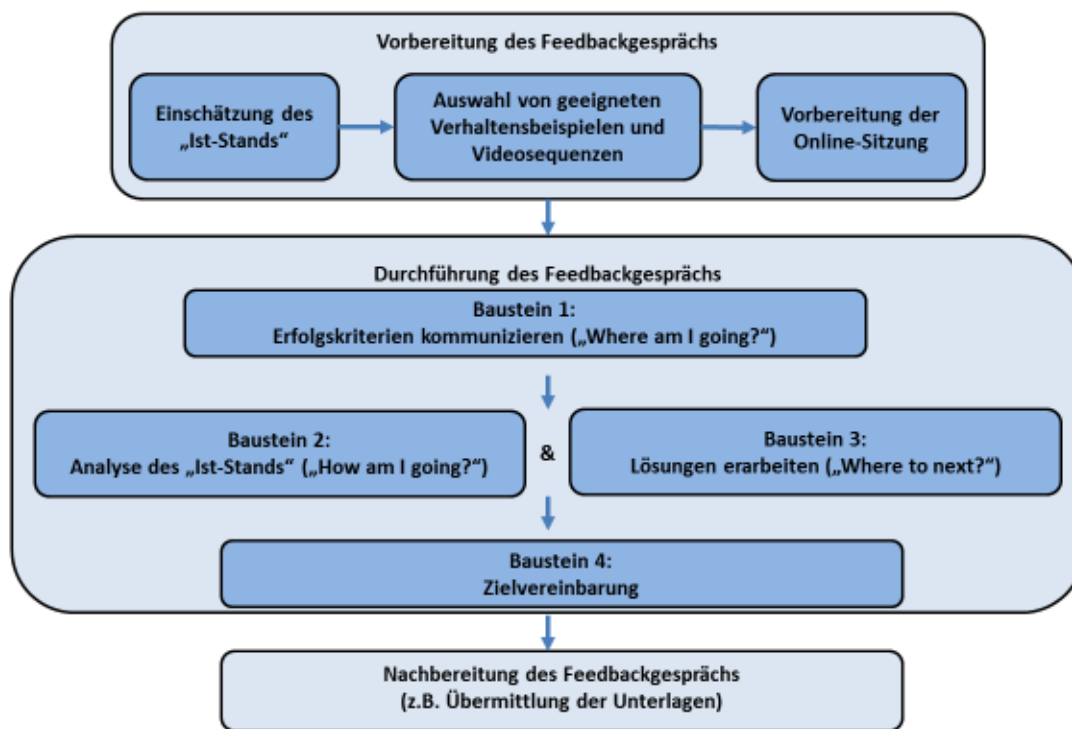


Abbildung 4: Konzept für ein online-basiertes Videofeedback

## 5.4 Erprobung des Feedback-Konzepts

Um sowohl die inhaltliche als auch die technische Umsetzung des Feedback-Konzepts zu erproben, wurde ein Pretest durchgeführt.

Dafür wurden online-basierte Feedbackgespräche mit Fachkräften durchgeführt, die am Projekt „BiSS-Evaluation“ teilgenommen haben. Im Rahmen des Projekts wurden bereits Videoaufnahmen von den Fachkräften erstellt, während sie in ihrem Alltag ein Bilderbuch mit drei- bis sechsjährigen Kindern anschauen. Es wurden alle Fachkräfte aus den teilneh-

menden Verbänden telefonisch kontaktiert, von denen eine Videoaufnahme vorlag ( $N = 14$ ). Die Teilnahme am Feedbackgespräch war freiwillig. Insgesamt nahmen vier Fachkräfte das Angebot an, ein individuelles Feedback bezüglich ihres sprachunterstützten Interaktionsverhaltens zu bekommen. Um möglichst verschiedene Blickwinkel zu berücksichtigen, wurden die vier Feedbackgespräche von der Autorin und zwei weiteren Projekt-Mitarbeiterinnen durchgeführt. Die beiden Mitarbeiterinnen waren bereits mit dem Konzept vertraut und erhielten zusätzlich eine ca. eintägige Schulung für die Durchführung des Feedbacks. Im Rahmen der Schulung wurden der Leitfaden und die Arbeitsvorlage schrittweise durchgesprochen. Zusätzlich wurde anhand kurzer Videosequenzen geübt, welches Feedback man auf Basis verschiedener Verhaltensweisen geben kann.

Nach dem Feedbackgespräch wurden die Fachkräfte gebeten, eine schriftliche Rückmeldung zum Feedback zu verfassen. Um ein möglichst wenig selektiertes und gefiltertes Feedback zu erhalten, wurden den Fachkräften folgende Fragen im offenen Antwortformat gestellt:

- Was haben Sie aus dem Feedbackgespräch für Ihre zukünftige Arbeit mitgenommen?
- Was hat Ihnen am Feedbackgespräch gut gefallen?
- Was hat Ihnen nicht so gut gefallen? Welche Anregungen haben Sie?
- Wie beurteilen Sie die technische Umsetzung und die Möglichkeiten eines online-basierten Feedbackgesprächs?

Zusätzlich fand ein mündlicher und schriftlicher Austausch über die Erfahrungen der Feedbackgeber statt. Die Erfahrungen der drei Feedbackgeber in den vier Gesprächen zeigten, dass das Feedback-Konzept prinzipiell gut umgesetzt werden konnte. Es war gut möglich anhand des Leitfadens und der Arbeitsvorlage das Gespräch vorzubereiten und durchzuführen. Es wurden kleinere Anpassungen in der Arbeitsvorlage vorgenommen, um die Handhabbarkeit noch zu verbessern. So wurden beispielsweise unnötige Informationen für eine bessere Übersichtlichkeit entfernt, Textfelder verkleinert bzw. vergrößert und am Anfang jeden Indikators Felder eingefügt, in denen der Name des zu besprechenden Indikators und die dazugehörigen Videosequenzen auf einen Blick zu erkennen sind. Die Fachkräfte haben das Feedbackgespräch durchweg sehr positiv bewertet. Als besonders hilfreich haben alle vier Fachkräfte die Videosequenzen genannt, da diese ihnen dabei geholfen haben, dem Feedbackgespräch zu folgen und die Gelegenheit gegeben haben, sich selbst mal aus einer anderen Perspektive zu betrachten. Die Gespräche hatten eine Dauer zwischen 60 und 90 Minuten. Die Fachkräfte gaben uns die Rückmeldung, dass das Gespräch idealerweise weniger als 90 Minuten dauern sollten, da es sonst schwer fällt, die

Aufmerksamkeit zu halten und das Gespräch im Alltag unterzubringen. Bezüglich der technischen Umsetzung ist positiv anzumerken, dass alle Feedbackgespräche durchgeführt werden konnten. Allerdings kam es bei allen vier Gesprächen zu technischen Problemen (z.B. in Form einer verzögerten Tonübertragung oder Unterbrechungen bei der Webcam-Übertragung). Dies konnte auf eine zu langsame Internetverbindung in der Kita zurückgeführt werden.

## 6 Umsetzung und Überprüfung des Feedback-Konzepts: Fragestellung und Methode

Die wissenschaftliche Überprüfung des Konzepts war Ziel der durchgeführten Studie. Die zugrunde liegenden Annahmen und Fragestellungen sowie die methodische Umsetzung der Untersuchung werden in diesem Kapitel thematisiert.

### 6.1 Fragestellungen und Hypothesen

Die Ergebnisse aus der Weiterbildungsforschung (z.B. Markussen-Brown et al., 2017; Werner et al., 2016) zeigen, dass das Interaktionsverhalten von frühpädagogischen Fachkräften durch die Teilnahme an Fortbildungen und Trainings optimiert werden kann. Bei den evaluierten Interventionen handelt es sich in der Regel um intensive bzw. langfristig angelegte Programme (z.B. „Heidelberger Interaktionstraining“ oder „MyTeachingPartner“). Als besonders wirksam haben sich dabei vor allem praxisnahe Formate erwiesen (Egert, 2015). Die positiven Effekte von Feedback, als Beispiel für ein solches Format, konnten in vielen Studien in klassischen Lehr-/Lernsettings bestätigt werden (z.B. Hattie & Timperley, 2007, Kluger & DeNisi, 1996). Handlungsleitend für die vorliegende wissenschaftliche Untersuchung war die Frage, ob ein einmaliges Feedbackgespräch im Sinne des dargestellten Konzepts ein ebenso wirksames und hilfreiches Instrument der Qualitätsentwicklung im frühpädagogischen Bereich sein kann.

Nach dem 4-Ebenen-Modell von Kirkpatrick (1998, S. 19) gibt es vor allem vier Kriterien für den Erfolg von Maßnahmen im Bildungsbereich: „reaction“ (Zufriedenheit), „learning“ (Lernerfolg), „behavior“ (Lerntransfer) und „results“ (Auswirkungen/Unternehmenserfolg).

Die Bedeutung der Zufriedenheit der Teilnehmer ist naheliegend, da sie – nicht zuletzt auch wegen der Motivation – ausschlaggebend dafür ist, ob ähnliche Fortbildungen in Zukunft wahrgenommen werden oder nicht. Daraus ergibt sich folgende erste Fragestellung für diese Studie:

➔ **Fragestellung 1:** Wie bewerten die teilnehmenden Fachkräfte das Feedbackgespräch?

Eine alleinige Beurteilung der Qualität einer Maßnahme durch eine Zufriedenheitsabfrage bei den Teilnehmerinnen und Teilnehmern reicht jedoch nicht aus. Aus diesem Grund besteht der Kern der Untersuchung darin, zu prüfen, wie sich das Feedbackgespräch auf das tatsächliche sprachliche Interaktionsverhalten der Fachkräfte auswirkt (Lerntransfer). Wie

bereits erwähnt, deuten die Ergebnisse verschiedener Meta-Analysen aus der Feedbackforschung sowie von Studien aus der Weiterbildungsforschung im pädagogischen Kontext insgesamt auf lernförderliche Effekte von Feedback hin. Es werden daher folgende Hypothesen aufgestellt:

- ➔ **Hypothese 1:** Die Teilnahme an einem online-basierten Videofeedbackgespräch hat einen signifikant positiven Einfluss auf das sprachliche Interaktionsverhalten von fröhpädagogischen Fachkräften während einer Bilderbuchbetrachtung.
- ➔ **Hypothese 2:** Die Effekte des Feedbackgesprächs bestehen auch nach Kontrolle von personenbezogenen und situativen Variablen (z.B. Qualifikation, Berufserfahrung, Anzahl der beteiligten Kinder).

Wie in Kapitel 3.2.4 ausgeführt, kann die Art und Weise, wie Feedback genutzt wird, von verschiedenen Merkmalen des Feedbackempfängers abhängen. Der direkte und implizite leistungssteigernde Effekt von positiven Selbstwirksamkeitserwartungen wurde vielfach empirisch nachgewiesen (z.B. Heslin & Latham, 2004; Karl et al., 1993). Es wird deshalb von folgender Annahme ausgegangen:

- ➔ **Hypothese 3:** Eine hohe auf Feedback bezogene Selbstwirksamkeitserwartung wirkt sich positiv auf die Effekte des Feedbacks aus.

In vielen Studien zeigte sich zudem ein signifikant positiver Zusammenhang zwischen Leistung und Lernorientierung, aber nicht zur Leistungsorientierung (Payne et al., 2007). Darauf aufbauend werden folgende Hypothesen formuliert:

- ➔ **Hypothese 4:** Eine höhere Lernzielorientierung geht mit stärkeren positiven Effekten des Feedbacks einher.
- ➔ **Hypothese 5:** Die Leistungsorientierung der Fachkräfte steht nicht im Zusammenhang mit den Effekten des Feedbacks.

Über die genannten Aspekte hinaus könnten auch andere Faktoren einen Einfluss auf die Wirksamkeit des Feedbacks haben, zu welchen es jedoch bislang kaum belastbare wissenschaftliche Ergebnisse aus der Feedbackforschung gibt. Es stellen sich demnach folgende Fragen:

- ➔ **Fragestellung 2:** Wie wirkt sich die Einstellung gegenüber Feedback auf die Effekte des Feedbackgesprächs aus?

- ➔ **Fragestellung 3:** Welche Effekte hat die Motivation der Fachkräfte auf die Feedbackwirkung?
- ➔ **Fragestellung 4:** Welchen Einfluss hat die Reflexionsfähigkeit der Fachkräfte auf die Feedbackeffekte?
- ➔ **Fragestellung 5:** Wie hängt die Bewertung der Fachkräfte hinsichtlich des Feedbackgesprächs und dessen Nutzen mit der Wirkung des Feedbacks zusammen?

Während der Erprobung des Feedback-Konzepts sind einzelne technische Probleme bei der Durchführung der online-basierten Feedbackgespräche aufgetreten. Daraus ergibt sich die Frage, inwiefern eventuell auftretende technische Schwierigkeiten die Effekte des Feedbacks beeinflussen.

- ➔ **Fragestellung 6:** Welche Auswirkungen hat das Auftreten von technischen Problemen während des Feedbackgesprächs auf dessen Effekte?

## 6.2 Design

Das Hauptziel der Evaluation war es, festzustellen, ob sich die Fachkräfte durch das Feedback in ihrem sprachunterstützten Interaktionsverhalten verbessern. Um dies herauszufinden, wurde das Interaktionsverhalten der teilnehmenden Fachkräfte vor und nach dem Feedback gemessen (Messwiederholung). Zudem muss sichergestellt werden, dass diese Veränderung vor allem durch das Feedback bedingt ist. Die unabhängige Variable „Feedback“ wurde daher so manipuliert, dass es eine Experimental- und eine Kontrollbedingung gibt, wobei jeweils andere Personen einer Bedingung zugeordnet wurden („between-subjects“ Design). In der vorliegenden Studie wurde demnach ein 2 x 2-faktorielles Design angewendet.

Um systematische Unterschiede zwischen den Gruppen möglichst gering zu halten und somit die interne Validität zu erhöhen, wurden keine natürlich vorgefundenen, sondern zufällig zusammengestellte Gruppen untersucht (siehe Kapitel 6.4). Mit dem Ziel, die Ergebnisse gut übertragbar und verallgemeinerbar zu halten (externe Validität), wurde die Untersuchung im Kita-Alltag durchgeführt.

Die Evaluation erfolgte dementsprechend in einem „2-Gruppen-Pretest-Posttest-Design“ in Form einer experimentellen Felduntersuchung.

## 6.3 Variablen und Instrumente

In den kommenden Abschnitten werden die einzelnen Variablen beschrieben, die im Rahmen der Untersuchung manipuliert bzw. gemessen wurden.

### 6.3.1 Abhängige Variablen

Wie bereits in Kapitel 6.1 beschrieben, soll mit der vorliegenden Studie geprüft werden, inwiefern das sprachliche Interaktionsverhalten von Fachkräften durch ein Feedback verbessert werden kann. Dabei wird zunächst thematisiert, wie die sprachliche Anregungsqualität erfasst wurde.

#### „Classroom Assessment Scoring System“ (CLASS)

Die sprachliche Anregung durch die Fachkraft wurde mithilfe des „Classroom Assessment Scoring System“ (CLASS Pre-K; Pianta, La Paro et al., 2008) bewertet. Die Dimension „Language Modeling“ ist eine von zehn Dimensionen zur Einschätzung der Interaktionsqualität in der Kita (vgl. Kapitel 2.1.2) und wird mithilfe von fünf Indikatoren operationalisiert (vgl. Kapitel 5.2). Die Ausprägung dieser Indikatoren wird anhand verschiedener Verhaltensweisen („Behavioral Marker“) beschrieben. In Tabelle 4 sind alle Indikatoren der Dimension „Language Modeling“ mit ihren jeweiligen „Behavioral Marker(n)“ dargestellt.

Tabelle 4: Dimension "Language Modeling", entnommen aus "Classroom Assessment Scoring System Manual, PRE-K" (CLASS Pre-K; S. 75; Pianta, La Paro et al., 2008; Teachstone Training, LLC, Charlottesville, VA; Reprinted by permission)

Indicator	Behavioral Marker
Frequent conversation	Back-and-forth exchanges
	Contingent responding
	Peer conversations
Open-ended questions	Questions require more than a one-word response
	Students respond
Repetition and extension	Repeats
	Extends/elaborates
Self- and parallel talk	Maps own actions with language
	Maps student action with language
Advanced language	Variety of words
	Connected to familiar words and/or ideas



Wie die einzelnen Aspekte auf den drei Qualitätsstufen „low“, „mid“ und „high“ ausgeprägt sind, wird im Manual des Instruments beschrieben (Pianta, La Paro et al., 2008). Im Folgenden wird für jeden Indikator eine hohe Qualitätsausprägung skizziert.

- ➔ „Frequent conversation“: Es finden häufig Gespräche zwischen der Fachkraft und den Kindern statt. Diese bieten ausreichend Gelegenheit für die Kinder, Sprache zu gebrauchen. Die Fachkraft hört den Kindern aktiv zu und knüpft mit ihren Äußerungen an die der Kinder an. Auf diese Weise entstehen auch längere bzw. intensivere Dialoge mit den Kindern. In den Gesprächen lässt die Fachkraft den Kindern ausreichend Raum für eigene Fragen und Kommentare. Phasen, die sehr stark von der Fachkraft kontrolliert werden, kommen nur selten vor. Es sind auch Gespräche zwischen den Kindern zu beobachten.
- ➔ „Open-ended questions“: Die Fachkraft stellt den Kindern viele offene Fragen. Diese Fragen erfordern Antworten mit mehr als einem Wort oder sogar in komplexen Sätzen. Häufig ist die Antwort auf diese Fragen unbekannt (z.B. „Was denkst du?“). Die Fragen können aber auch auf eindeutige bekannte Antworten abzielen (z.B. „Was ist auf dieser Seite passiert?“). Dabei ist es nicht nur wichtig, dass die Fachkraft den Kindern offene Fragen stellt, sondern auch, dass Sie den Kindern die Möglichkeit gibt, darauf zu antworten (z.B. durch Abwarten und aufmerksames Zuhören).
- ➔ „Repetition and extension“: Die Fachkraft greift den Großteil der kindlichen Äußerungen auf und erweitert diese gegebenenfalls. Auf diese Art und Weise wird die Antwort des Kindes und deren Empfang bestätigt (z.B.: Fachkraft: „Was ist das?“, Kind: „Ein Baum.“, Fachkraft: „Genau- ein Baum.“). Die Äußerungen können zudem durch Hinzufügen zusätzlicher Informationen erweitert und/oder in eine komplexere Satzstruktur gebracht werden (z.B.: Kind: „Schau - Tier“, Fachkraft: „Das ist ein Tier. Man nennt es Leopard.“).
- ➔ „Self- and parallel talk“: Die Fachkraft begleitet die eigenen Handlungen und die der Kinder sprachlich. Das heißt, sie sagt den Kindern, was sie gerade macht (z.B. „Ich hole schnell das Buch.“) und beschreibt das, was die Kinder machen (z.B. „Oh- du zeigst uns alle Schmetterlinge auf der Seite.“).
- ➔ „Advanced language“: Die Fachkraft setzt einen reichhaltigen Wortschatz ein. Dabei wird auf verschiedene Wortarten in Form von Substantiven, Adjektiven, Adverbien oder Präpositionen zurückgegriffen. Die Fachkraft versucht den Wortschatz der Kinder zu erweitern, indem sie neue Begriffe einbringt (z.B.: Kind zählt Farben einer Blume auf und Fachkraft antwortet: „Das sind aber viele verschiedene Farben. Es ist eine sehr bunte Blume.“). Dabei baut sie auf Wissen und Konzepten auf, über welche(s) die Kinder bereits verfügen.

Anhand der Ausprägung dieser Indikatoren wird ein Wert zwischen 1 und 7 für den Aspekt „Language Modeling“ vergeben. Wenn das Verhalten gar nicht beobachtet wird, liegt die Qualität auf Stufe 1. Eine mittlere Bewertung (4) bedeutet, dass die Indikatoren zum Teil gut, aber nicht konsequent umgesetzt werden. Den höchsten Wert (7) erhält eine Fachkraft, wenn sie die genannten Verhaltensweisen mit einer hohen Intensität und Häufigkeit zeigt und davon (fast) alle Kinder profitieren. Die Werte 2, 3, 5 und 6 bilden jeweils die Übergänge zum nächst höheren oder niedrigeren Qualitätsbereich. In amerikanischen Studien erreichen Fachkräfte durchschnittlich in der Dimension „Language Modeling“ einen Wert von  $M = 2.85$  ( $SD = 0.73$ )<sup>3</sup> (Pianta, La Paro et al., 2008). In deutschen Kindergärten zeigen frühpädagogische Fachkräfte ein ähnliches Niveau sprachlicher Anregung. Suchodoletz et al. (2014) berichten beispielsweise eine durchschnittliche Sprachanregung von 2.73 ( $SD = 0.76$ ).

Die Inter-Rater-Übereinstimmung der CLASS wurde unter anderem anhand von Videoaufnahmen getestet, die von Fachkräften aufgenommen wurden, welche am Fortbildungsprogramm „MyTeachingPartner“ teilgenommen haben. Zwei Rater haben insgesamt 33 Videoaufnahmen mit einer Länge von 30 Minuten mit der CLASS bewertet. Als reliabel gilt eine Übereinstimmung von  $\pm 1$  auf der sieben-stufigen Skala. Für die Dimension „Language Modeling“ konnte bei 84.8 % der Videoaufnahmen eine solche Übereinstimmung festgestellt werden (Pianta, La Paro et al., 2008). Hinsichtlich der Validität der CLASS konnten verschiedene Studien signifikante Zusammenhänge mit anderen Instrumenten und kindlichen Entwicklungsmaßen feststellen (siehe Kapitel 2.1.2 und Kapitel 2.1.3).

Da die Bewertung der Dimension „Language Modeling“ die verschiedenen Indikatoren der sprachlichen Unterstützung in einem Wert zusammenfasst, ist die Einschätzung eher global und erlaubt keine Aussagen über Veränderungen in einzelnen Verhaltensaspekten (Mayer & Beckh, 2018). Um auch Verbesserungen auf Verhaltensebene besser nachweisen zu können und um ein genaueres Verständnis von den auftretenden Effekten zu bekommen, wurde ein zusätzliches Kodier-Schema für die Beurteilung des sprachlichen Interaktionsverhaltens eingesetzt.

### Kodier-Schema zur Einschätzung des sprachlichen Interaktionsverhaltens

Für eine kleinschrittigere Einschätzung des Interaktionsverhaltens wurde ein im Projekt „BiSS-Evaluation“ entstandenes Kodier-Schema (Radan, Schauland & Wirts, 2017) angepasst und erweitert. Als Grundlage diente die Dimension „Language Modeling“ aus der CLASS. Auf der Basis von Literatur zur sprachlichen Entwicklung und Förderung von Kindern im Kindergartenalter wurden die Aspekte der CLASS erweitert und ergänzt.

---

<sup>3</sup>  $M$  = Mittelwert;  $SD$  = Standardabweichung

Hinsichtlich der durch die Fachkraft gestellten Fragen wurde eine genauere Differenzierung vorgenommen, welche sich an die Operationalisierung von Siraj-Blatchford und Manni (2008) anlehnt. Im Rahmen der Studie „Researching Effective Pedagogy in the Early Years“ (REPEY), welche auf Daten des EPPE-Projekts aufbaut, wurden besonders effektive pädagogische Strategien von erfolgreichen Kindergarten-Settings identifiziert. In einer erweiterten Analyse haben Siraj-Blatchford und Manni (2008) 5808 Fragen aus 400 Stunden Beobachtung ausgewertet. Sie konnten die Fragen in sieben verschiedene Arten von geschlossenen Fragen und vier Formen von offenen Fragen einordnen. Tabelle 5 zeigt die verschiedenen Fragetypen von offenen und geschlossenen Fragen.

Tabelle 5: Kategorisierung von Fragetypen nach Siraj-Blatchford und Manni (2008)

Offene Fragetypen	Geschlossene Fragetypen
<b>Open question:</b> mehr als eine richtige/akzeptierte Antwort; nicht unbedingt Erwartungen bzgl. der Antwort	<b>Closed question with one known:</b> Fachkraft kennt die Antwort und es gibt nur eine oder eine begrenzte Anzahl an richtigen/akzeptierten Antworten
<b>Apparently open:</b> scheint eine offene Frage zu sein, aber in den Daten keine kindliche Antwort (z.B. wegen Unterbrechung)	<b>Closed question answer not known:</b> Antwort ist unbekannt/offen, aber typischerweise kurz; auch Entscheidungsfragen
<b>Pseudo Wonder:</b> richtige Antwort ist vllt. bekannt, aber Kind kann frei antworten, nicht nur wiedergeben	<b>Yes-No question:</b> Frage erfordert Ja-/Nein-Antwort
<b>Tag question:</b> um kindliche Aktivität/ Beteiligung im Spiel aufrechtzuerhalten oder anzuregen	<b>Pseudo-choice question:</b> indirekte oder vermeintliche Fragen: täuschen eine Wahlmöglichkeit vor; Aussage oder Aufforderung, die als Frage formuliert ist
	<b>Question that informs about rules:</b> Frage dient dazu, Verhalten der Kinder zu steuern
	<b>Statement question:</b> Aussagen, Beobachtungen, Fakten oder Lob, die als Frage formuliert sind
	<b>Recall elicit question:</b> Erinnerungsfragen zum Abruf von bereits gelerntem oder erfahrenem Wissen

Diese Kategorisierung wurde als Grundlage für die Kodierung der Fragen in der vorliegenden Studie verwendet, wobei verschiedene Anpassungen vorgenommen wurden. Hinsichtlich einer möglichen Antwort wurde nicht danach unterschieden, ob diese der Fachkraft bekannt oder unbekannt ist. In diesem Projekt steht außerdem das Anregungspotenzial des Verhaltens der Fachkraft im Mittelpunkt. Es wurde deshalb nicht berücksichtigt, ob die Kinder antworten oder nicht. Die Kategorie „Tag Question“ wurde nicht übernommen, da hier

nur der Fokus auf sprachlicher Anregung liegt. Die Fragetypen „Pseudo-choice question“ und „Question that informs about rules“ wurden zusammengefasst, da es sich bei Beiden um Fragen ohne Anregungscharakter handelt. Die Kategorie „Statement question“ wurde mit zu „Yes-No question“ gezählt, da auch auf diese nur mit „Ja“ oder „Nein“ geantwortet werden kann. Sogenannte „Recall elicit questions“ wurden nicht als extra Fragentyp aufgenommen, da es nicht um kognitive, sondern um sprachliche Anregung geht. Eine Sichtung der Daten aus dem Projekt „BiSS-Evaluation“ zeigte, dass Fachkräfte zum Teil zwar Fragen stellen, den Kindern aber formal keine Möglichkeit zum Antworten gegeben wird (z.B. indem die Antwort selbst formuliert wird oder direkt eine andere Frage gestellt wird). Aus diesem Grund wurden die einzelnen Fragentypen jeweils differenziert in „mit Antwortmöglichkeit“ und „ohne Antwortmöglichkeit“.

Da Kinder nicht nur durch Fragen zum Sprechen angeregt werden können, wurden sogenannte Erzählaufforderungen in das Kodier-Schema einbezogen. Auch hier wurde die Antwortmöglichkeit der Kinder berücksichtigt.

Bei der Wiederholung und Erweiterung der kindlichen Äußerungen wurde zudem in Anlehnung an die Literatur (z.B. Reichert-Garschhammer & Kieferle, 2011) die korrektive Funktion mit einbezogen. Sowohl Wiederholungen als auch Erweiterungen werden daher in „korrektiv“ und „nicht korrektiv“ unterschieden.

In Bezug auf die Gespräche bei der Bilderbuchbetrachtung wurden zwei Aspekte in das Kodier-Schema aufgenommen: „Längere Gespräche“ und „Anknüpfen“. Der Kode „Längere Gespräche“ bezieht sich auf den Aspekt „Back-and-forth exchanges“ aus der CLASS. Als längeres Gespräch werden Dialoge mit einem Kind bewertet, die aus mindestens vier Sprecherwechseln bestehen<sup>4</sup>. Dabei ist es wichtig, dass die Äußerungen sich aufeinander beziehen. Dies können zum Beispiel Fragen an das Kind sein. Ein Gespräch mit einem Kind kann auch aufgebaut werden, indem man an die Äußerung des Kindes mit einer eigenen Aussage anknüpft, da diese wiederum für das Kind Möglichkeiten der Erwiderung beinhaltet. Der Aspekt „Anknüpfen“ korrespondiert mit dem „Contingent responding“ in der CLASS.

Die auf theoretischer Basis entwickelten Codes wurden an insgesamt 13 Videoaufnahmen von Bilderbuchbetrachtungen getestet. Auf der Grundlage dieser Auswertungen wurde die Beschreibung der Codes zum Teil noch angepasst und differenzierter formuliert und Beispiele für alle Codes hinzugefügt. In Tabelle 6 sind alle Codes sowie deren inhaltliche Beschreibung und jeweils passende Beispiele zusammengefasst. Das komplette Kodier-Schema findet sich in Anhang B.

---

<sup>4</sup> Eine ähnliche Vorgabe findet sich zum Beispiel bei Milburn et al. (2014) mit fünf Sprecherwechseln.

Tabelle 6: Kodierung des sprachlichen Interaktionsverhaltens

Kode	Beschreibung	Beispiel
Längere Gespräche	Mind. 4-5 Sprecherwechsel mit einem Kind; Äußerungen müssen Bezug aufeinander nehmen, sich auf ein Thema beziehen und eine inhaltliche Aussage enthalten	
Geschlossene Fragen mit Antwortmöglichkeit	<p>Arten von geschlossenen Fragen:</p> <ul style="list-style-type: none"> <li>- Fachkraft kennt die Antwort und es gibt nur eine oder eine begrenzte Anzahl an richtigen/akzeptierten Antworten</li> <li>- die Antwort ist nicht bekannt/offen, aber die Frage erfordert nur eine kurze Antwort</li> <li>- Entscheidungsfragen</li> </ul>	<p>Was ist das?</p> <p>Was ist deine Lieblingsfarbe?</p> <p>Magst du lieber Hunde oder Katzen?</p>
Geschlossene Fragen ohne Antwortmöglichkeit	Das Kind hat keine Möglichkeit, zu antworten (z.B. wenn Antwort selbst formuliert wird oder direkt eine andere Frage oder Erzählaufforderung folgt.	
Ja/Nein-Fragen mit Antwortmöglichkeit	<p>Fragen, auf die das Kind nur mit Ja oder Nein antworten kann:</p> <ul style="list-style-type: none"> <li>- Ja/Nein- Fragen</li> <li>- „Bestätigende“ Fragen: Äußerung, auf die eine bestätigende Antwort erwartet wird.</li> <li>- Nachhaken bei falschen Antworten</li> </ul>	<p>Ist der Junge traurig?</p> <p>Das ist ein großer Baum, oder?</p> <p>Bist du sicher, dass...?</p>
Ja/Nein-Fragen ohne Antwortmöglichkeit	Siehe oben	
Vermeintliche Fragen	<p>Aussagen oder Aufforderungen, die als Frage formuliert, aber nicht als Frage gemeint sind- das heißt es wird nicht wirklich eine Antwort vom Kind erwartet.</p> <ul style="list-style-type: none"> <li>- Indirekte Verhaltensaufforderungen</li> <li>- Fragen, mit denen Kinder an Regeln erinnert werden.</li> <li>- Rhetorische Fragen</li> </ul>	<p>Kannst du dich bitte dort hinsetzen?</p> <p>Siehst du wie schön Max sitzt?</p> <p>Hab ich es euch nicht gesagt?</p>
Offene Fragen mit Antwortmöglichkeit	<p>Fragen regen Antworten von mehr als einem Wort an (nicht Zwei-Wort-Äußerungen mit Artikel oder Präposition)</p> <ul style="list-style-type: none"> <li>- es gibt mehr als eine richtige/ akzeptierte Antwort</li> </ul>	Warum hat der Hase Angst?

## Methode

	<p>- richtige Antwort ist evtl. bekannt, aber das Kind kann frei antworten</p> <p>Bei kombinierten Äußerungen zählt es als offene Frage, wenn am Ende der Äußerung ein Fragezeichen stehen würde.</p>	<p>Was passiert auf dieser Seite?</p> <p>Erzähl du Daniel- was isst du gerne?</p>
Offene Fragen ohne Antwortmöglichkeit	Siehe oben	
Aufforderungen zum Sprechen mit Antwortmöglichkeit	<p>Aufforderungen/Ermutigungen zum Sprechen, die keine Frageform aufweisen; Können auch indirekte Aufforderungen sein (z.B. durch Sprechpause und Blickkontakt)</p> <p>Bei kombinierten Äußerungen zählt es als Erzählaufforderung, wenn am Ende der Äußerung ein Punkt oder Ausrufezeichen stehen würde.</p>	<p>Erzählt mir, was ihr auf dem Bild seht.</p> <p>Was ist das? Ein...</p> <p>Was isst du gerne- erzähl doch mal!</p>
Aufforderungen zum Sprechen ohne Antwortmöglichkeit	Siehe oben	
Selftalk	Versprachlichen eigener Handlungen (wenn Fachkraft parallel die Handlung ausführt).	Ich blättere mal die Seite um.
Paralleltalk	<p>Versprachlichen kindlicher Handlungen</p> <p>NICHT: in die Zukunft gerichtete Aussagen oder Aufforderungen zum Handeln (z.B. „Jetzt zählen wir mal die Affen auf dem Bild.“)</p>	Oh, du zeigst uns alle Schmetterlinge auf dem Bild.
Wiederholung ohne Korrektiv	Reine Wiederholung der kindlichen Äußerung, ohne Umformung oder Korrektur	K: 2 Katzen, FK: (Ja-) 2 Katzen.
Korrektive Wiederholung	Wiederholung der kindlichen Äußerung, wobei „Fehler“ reaktiv aufgenommen werden (z.B. z.B. falscher/fehlender Artikel, inhaltlich falsche Antwort, falsch gebeugtes Verb)	K: Die Baum, FK: Der Baum.
Korrektive Erweiterung	<p>„Fehlerhafte“ kindliche Äußerungen werden aufgenommen und erweitert:</p> <p>- auf semantischer und/oder syntaktisch-morphologischer Ebene</p>	K: Die Baum, FK: Der große Baum.
Erweiterung ohne Korrektiv	<p>Äußerungen werden wiederholt und erweitert:</p> <p>- auf semantischer und/oder syntaktisch-morphologischer Ebene</p>	<p>K: Ein Baum, FK: Ja- das ist ein Baum.</p> <p>K: Ein Baum. FK: Ein Baum. Das ist eine Eiche.</p>
Anknüpfen	Fachkraft reagiert auf eine Äußerung des Kindes (nur wenn keiner der anderen Codes zutrifft)	K: Was ist das? FK: Das ist eine Zange.

Das Vorkommen der einzelnen Aspekte während der Bilderbuchbetrachtung wurde gezählt. Für jede Bilderbuchbetrachtung wurde also die Anzahl der sprachlichen Verhaltensweisen ermittelt. Für jede Äußerung der Fachkraft konnte nur ein Kode vergeben werden. Wenn bei einer Äußerung zwei Kodes zutreffen (z.B. eine Wiederholung, die als Frage intoniert ist), wurde immer der höherwertige Kode vergeben. Für diesen Fall wurde aufgrund des vermuteten Anregungspotenzials folgende Rangfolge erstellt:

1. Offene Fragen
2. Erzählaufforderungen
3. Geschlossene Fragen
4. a) Korrektive Erweiterung  
b) Korrektive Wiederholung
5. Erweiterung ohne Korrektiv
6. a) Wiederholung ohne Korrektiv  
b) Ja-/Nein-Fragen
7. Vermeintliche Fragen
8. Anknüpfen

### Einschätzung der sprachlichen Anregung

Für die Beurteilung der sprachlichen Unterstützung wurde in der vorliegenden Untersuchung von jeder teilnehmenden Fachkraft zu T1 und zu T2 jeweils eine Videoaufnahme während einer Bilderbuchbetrachtung im Kita-Alltag gemacht. Voraussetzung war, dass die Kinder zwischen drei und sechs Jahren alt waren. Zudem musste sichergestellt werden, dass die Situationen vergleichbar sind. Die Fachkräfte wurden deshalb gebeten, bei T1 und T2 das gleiche oder ein sehr ähnliches Bilderbuch mit möglichst den gleichen Kindern anzuschauen. Darüber hinaus wurden keine Einschränkungen oder Vorgaben gemacht, um die Situation so nah wie möglich am Alltag zu belassen und so wenig wie möglich zu verzerren.

Die Videoaufnahmen wurden von vier Ratern (einschließlich der Autorin) ausgewertet. Alle Rater sind für die Anwendung der CLASS Pre-K zertifiziert. Für die Zertifizierung ist die Teilnahme an einem mehrtägigen Training sowie eine jährliche Überprüfung der Reliabilität<sup>5</sup> notwendig. Die Rater haben zudem durch zahlreiche CLASS-Erhebungen im Rahmen des Projekts „BiSS-Evaluation“ umfangreiche Erfahrung in der Verhaltensbeobachtung gesammelt. Zusätzlich haben die Rater vor der eigentlichen Auswertung insgesamt 13 Videoaufnahmen mithilfe der CLASS und des neu entwickelten Kodier-Schemas eingeschätzt. Bei den ersten Videoaufnahmen zeigten sich zum Teil Probleme bei der Bewertung mithilfe des

---

<sup>5</sup> Um die Zertifizierung zu behalten, müssen jedes Jahr fünf Videos ausgewertet und dabei eine Übereinstimmung von 80 % mit einem Expertenrating erreicht werden.

Kodier-Schemas. Die Ergebnisse der Auswertungen wurden daher in regelmäßigen Abständen ausführlich diskutiert. Auf diese Weise fand eine Art Kalibrierung der Rater statt. Ebenso konnte diese Testphase für eine Verbesserung des Kodier-Schemas genutzt werden.

Die Auswertung erfolgte blind, das heißt die Rater wussten nicht, aus welcher Bedingung oder von welchem Messzeitpunkt die entsprechende Videoaufnahme stammt. Da die Rater zum Teil selbst Videoaufnahmen angefertigt haben, wurde sichergestellt, dass ein Rater niemals eine Videoaufnahme von einer Kita ausgewertet hat, die selbst besucht wurde. Um sicherzustellen, dass die Rater nicht (un)bewusst nach Verhaltensunterschieden zwischen den Videoaufnahmen einer Fachkraft suchen, wäre es hilfreich, wenn die beiden Videoaufnahmen von einer Fachkraft (T1 und T2) immer von unterschiedlichen Ratern ausgewertet werden. Wie bereits thematisiert wurde, wird eine Bewertung mit der CLASS jedoch als reliabel angesehen, wenn eine Übereinstimmung von  $\pm 1$  vorliegt. Wenn die Videoaufnahmen einer Fachkraft zu T1 und T2 von unterschiedlichen Ratern eingeschätzt werden, könnte dies demnach dazu führen, dass mögliche kleine Effekte über- bzw. unterschätzt werden. Da die Rater die Videos weder einer Bedingung noch einem Zeitpunkt zuordnen konnten, wurde die Auswertung der Videoaufnahmen von T1 und T2 in der vorliegenden Studie immer vom gleichen Rater vorgenommen. Um zu vermeiden, dass die Videoaufnahmen dennoch miteinander verglichen werden können, wurde zudem ein möglichst großer Zeitabstand zwischen die Auswertung von T1 und T2 gelegt. Bei der Auswertung wurde nur die Bilderbuchbetrachtung bewertet. Andere Handlungen oder Aktivitäten, die auf der Videoaufnahme zu sehen sind, wurden nicht berücksichtigt.

Für jede Videoaufnahme gibt es dementsprechend einen Wert für die CLASS-Dimension „Language Modeling“ sowie die Anzahl der einzelnen Verhaltensweisen, welche mit dem neu entwickelten Kodier-Schema bestimmt wurde (z.B. Anzahl offener Fragen mit Antwortmöglichkeit). Die Häufigkeitsangaben zu den Verhaltensweisen beziehen sich dabei auf die gesamte Bilderbuchbetrachtung einer Fachkraft. In einigen Studien (z.B. Simon & Sachse, 2011) wird nur ein Teil der Videoaufnahme, z.B. die ersten fünf Minuten, analysiert, um den Bezugszeitraum für die Auswertung gleichzuhalten. In der vorliegenden Studie wurde bewusst die gesamte Bilderbuchbetrachtung ausgewertet. Auch Pentimonti et al. (2017) beziehen sich für die Erfassung der Häufigkeit von „Scaffolding“-Strategien in der Kita auf das gesamte Material. Für die hier durchgeführte Studie kann dieses Vorgehen damit begründet werden, dass eine Bilderbuchbetrachtung sehr unterschiedlich gestaltet werden kann und somit eine Auswertung eines kleinen Ausschnitts nicht unbedingt allen Fachkräften gerecht wird. Manche Fachkräfte lesen zum Beispiel zunächst aus dem Buch vor, um die Aufmerksamkeit der Kinder zu gewinnen und beziehen die Kinder erst später ein. Das Ziel



des Feedbacks bestand zudem nicht darin, die Fachkräfte zu instruieren, so viele sprachförderliche Strategien wie möglich in kurzer Zeit umzusetzen, sondern dass sie die Bilderbuchbetrachtung möglichst sprachförderlich für die Kinder gestalten. Für die Beurteilung der sprachlichen Anregung und der Effekte des Feedbacks ist daher die gesamte Situation relevant. Aus diesem Grund erfolgte auch keine Auswertung der sprachlichen Anregung pro Minute. So wäre beispielsweise folgender Fall denkbar: Eine Fachkraft stellt zu T1 fünf offene Fragen, wobei die Bilderbuchbetrachtung fünf Minuten dauert. Zu T2 werden nun 10 offene Fragen an die Kinder gestellt und die Bilderbuchbetrachtung hat eine Dauer von 10 Minuten. Betrachtet man die Anzahl an offenen Fragen pro Minute, kann nicht von einer Verbesserung ausgegangen werden. Für die sprachliche Anregung der Kinder kann dies jedoch einen bedeutenden Unterschied ausmachen.

Zur Überprüfung der Reliabilität fand eine Doppelkodierung bei einem Teil der Videoaufnahmen statt. Das Vorgehen kann in Kapitel 6.6.1 nachvollzogen werden. Die Ergebnisse werden in Kapitel 7.3.1 beschrieben.

### Bewertung des Feedbacks durch die teilnehmenden Fachkräfte

Neben den Effekten auf Performanz-Ebene stellt die Einschätzung der teilnehmenden Fachkräfte ein wichtiges Maß für den Erfolg des Feedbacks dar.

Diese wurde unter anderem mit dem „Questionnaire for Professional Training Evaluation“ (Q4TE) von Grohmann und Kauffeld (2013) erhoben. Der Fragebogen wurde zur Evaluation von Trainings entwickelt und so formuliert, dass er in verschiedenen Kontexten einsetzbar ist. Im Fragebogen werden die vier Ebenen im Modell von Kirkpatrick (1998) aus der Perspektive der Fachkräfte bewertet. Die psychometrischen Eigenschaften des Fragebogens wurden in drei verschiedenen Studien mit insgesamt 1134 Angestellten untersucht. Die Überprüfung der faktoriellen Struktur weist auf eine Lösung mit sechs Faktoren und jeweils zwei Items pro Faktor hin. Die sechs Faktoren werden den vier Ebenen „Reaction“, „Learning“, „Behavior“ und „Organizational results“ aus dem 4-Ebenen-Modell von Kirkpatrick (1998) zugeordnet. Auf einer elf-stufigen Skala wird bewertet, inwiefern man den Aussagen der einzelnen Items zustimmt. Die Faktoren auf den vier Ebenen mit den entsprechenden Werten für die interne Konsistenz werden in Tabelle 7 dargestellt.

Tabelle 7: Skalen des „Questionnaire for Professional Training Evaluation“ von Grohmann und Kauffeld (2013) sowie die dazugehörigen Cronbachs Alpha-Werte

	Skala	Inhalt	Cronbachs $\alpha$ (Autoren)	Cronbachs $\alpha$ (vorl. Studie)
Reaction	Satisfaction	Gute Erinnerung Spaß	0.91	0.93
	Utility	Bringt viel Nützlich für Arbeit	0.96	0.93
Learning	Knowledge	Mehr Wissen Neues gelernt	0.91	0.88
Behavior	Application to practice	Kenntnisse nutzen Inhalte anwenden	0.90	0.83
Organizational results	Individual	Zufriedenheit mit Arbeit Verbesserung der Arbeits- leistung	0.86	0.92
	Global	Arbeitsabläufe vereinfacht Unternehmensklima ver- bessert.	0.79	

Anmerkung: Die Ermittlung der Cronbachs Alpha-Werte basiert auf einer Stichprobengröße von  $N = 40$  (Satisfaction und Utility) bzw.  $N = 39$  (Knowledge, Application to practice und Individual organizational results).

Für die vorliegende Studie wurden die Originalitems auf den Kontext dieser Studie angepasst. Das Wort „Training“ wurde durch „Feedback“ oder „Feedbackgespräch“ ersetzt. Anstatt „Arbeit“ oder „Arbeitsleistung“ wurde zum Teil der Begriff „Arbeitsweise“ verwendet. Die Skala „Global Organizational results“ wurde nicht eingesetzt, da das Feedbackgespräch ausschließlich auf das Verhalten der Fachkraft fokussiert und wenig mit dem Klima oder anderen Arbeitsabläufen in der Kita zu tun hat. Die Beantwortung der Items erfolgte in anderer Reihenfolge in einem schriftlichen Fragebogen zu T2. Die für die vorliegende Studie berechneten Cronbachs Alpha-Werte entsprechen in etwa den von den Autoren des Instruments berichteten Werten. Da die Werte für alle Skalen über 0.80 liegen, kann die innere Konsistenz als ausreichend hoch betrachtet werden<sup>6</sup>. Die Items der verschiedenen Skalen wurden zu einer Variable zusammengefasst, welche jeweils den Mittelwert beider Items darstellt.

<sup>6</sup> Nach Döring und Bortz (2016) wird bei allen Werten  $> 0.80$  von einer guten inneren Konsistenz ausgegangen.

Um einen noch tieferen Einblick zu bekommen, wurde zusätzlich zum Q4TE auch mit einem offenen Antwortformat die Einschätzung der Fachkräfte bezüglich des Feedbackgesprächs erfasst. Die Fachkräfte wurden in dem schriftlichen Fragebogen zu T2 gebeten, folgende Fragen zu beantworten:

- Was haben Sie aus dem Feedbackgespräch für Ihre pädagogische Arbeit (bei der Bilderbuchbetrachtung) mitgenommen?
- Was haben Sie davon nach dem Feedbackgespräch bereits umgesetzt?
- Was hat Ihnen am Feedbackgespräch gut gefallen?
- Was hat Ihnen nicht so gut gefallen? Welche Anregungen haben Sie?
- Wie beurteilen Sie die technische Umsetzung und die Möglichkeiten eines online-basierten Feedbackgesprächs?

Die Antworten der Fachkräfte wurden systematisiert und zusammengefasst.

### 6.3.2 Intervention

Die unabhängige Variable „Feedback“ wurde so manipuliert, dass zwei Bedingungen hergestellt wurden: Experimental- und Kontrollbedingung. In der Experimentalbedingung haben die Fachkräfte an einem online-basierten Feedbackgespräch teilgenommen. Die Fachkräfte in der Kontrollbedingung erhielten kein online-basiertes Feedback.

Die Feedbackgespräche wurden nach dem in Kapitel 5 vorgestellten Feedback-Konzept durchgeführt. Im Feedbackgespräch wurde das sprachanregende Interaktionsverhalten der Fachkraft anhand der Videoaufnahme von T1 reflektiert. Die Vorbereitung und Durchführung des Feedbackgesprächs erfolgte anhand der bereits erwähnten Arbeitsvorlage. Die ausgewählten Videosequenzen wurden mithilfe der Software „PowerDirector“ geschnitten. Die notwendige Technik für das Feedbackgespräch (Webcam und Headset) wurde der Fachkraft zur Verfügung gestellt. Die Fachkräfte haben am Feedbackgespräch über einen Computer in einem Raum der Kita teilgenommen, in dem sie für den Zeitraum des Gesprächs ungestört waren. Die Feedbackgespräche wurden von der Autorin sowie einer Mitarbeiterin des Staatsinstituts für Frühpädagogik durchgeführt. Die Mitarbeiterin hatte auch für den Pretest Feedbackgespräche moderiert und war deshalb bereits geschult.

Der Ablauf eines Feedbackgesprächs soll mit den folgenden Abbildungen veranschaulicht werden. Dafür wurden Screenshots von einer nachgestellten Feedbacksitzung angefertigt. Die Abbildung 5 zeigt den „virtual team room“ von vitero. Im oberen Bereich sind die Teilnehmer der Sitzung zu sehen.

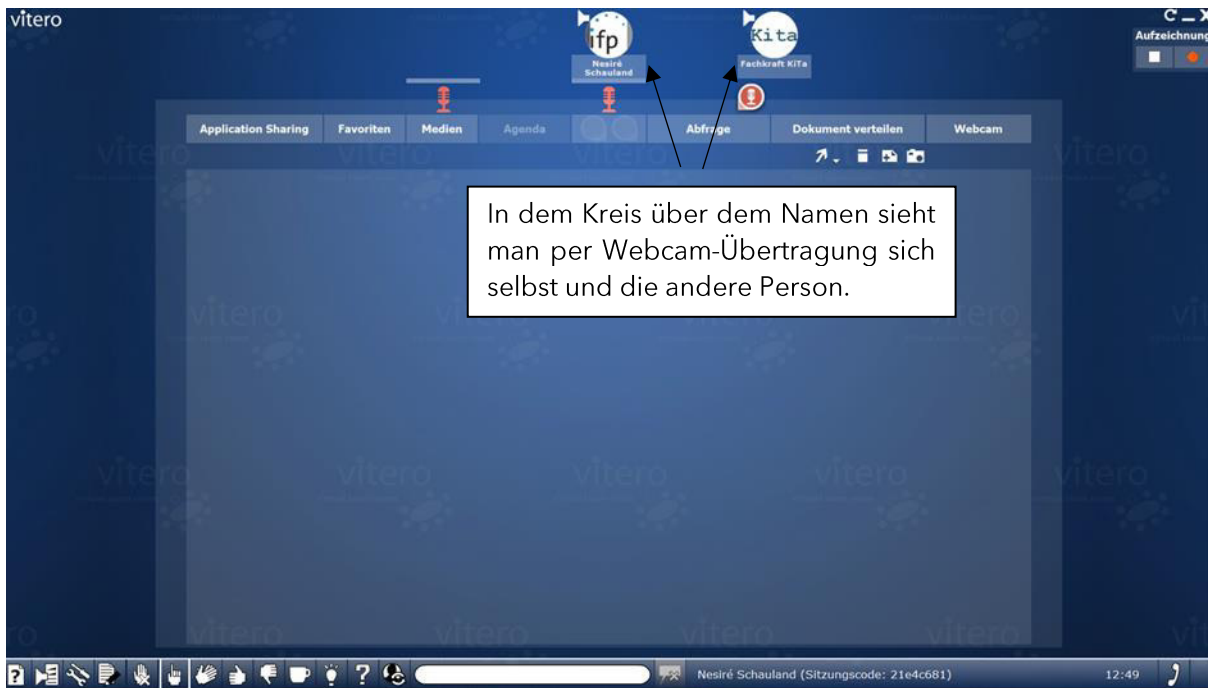


Abbildung 5: Ansicht vom "virtual team room" beim Feedbackgespräch

Das Feedback hat mit einer Darstellung der Indikatoren für sprachförderliches Verhalten begonnen. Zur Veranschaulichung wurden der Fachkraft die Indikatoren und die dazugehörigen „Behavioral Marker“ im „virtual team room“ angezeigt. Der Feedbackgeber kann dazu über das Dialogfeld „Medien“ das entsprechende Dokument aufrufen, welches dann in der freien Fläche in der Mitte („Tisch“) seitenweise angezeigt wird (siehe Abbildung 6). Vitero kann dabei auf Dateien zurückgreifen, die zuvor auf dem Computer des Moderators (Feedbackgeber) gespeichert wurden.

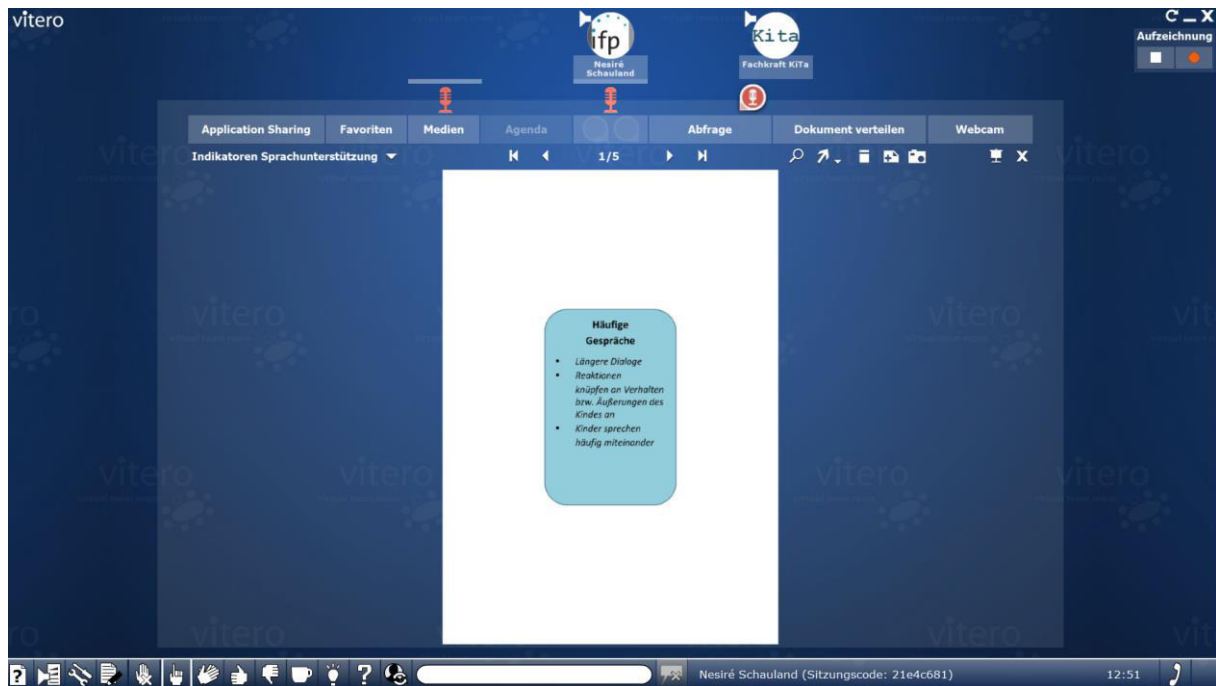


Abbildung 6: Baustein 1: Erfolgskriterien kommunizieren („Where am I going?“)

Im Anschluss an die Darstellung der Indikatoren wurden die einzelnen Videosequenzen mit der Fachkraft reflektiert. Diese können ebenfalls über das Menü „Medien“ abgerufen und in der Mitte des „Tisches“ angezeigt werden (siehe Abbildung 7). Nur der Moderator (Feedbackgeber) kann die einzelnen Videosequenzen starten bzw. stoppen. Die Fachkraft kann entweder „Stopp“ sagen oder mithilfe der Melde-Funktion (erhobener Zeigefinger) dem Feedbackgeber signalisieren, dass die Aufnahme angehalten werden soll.

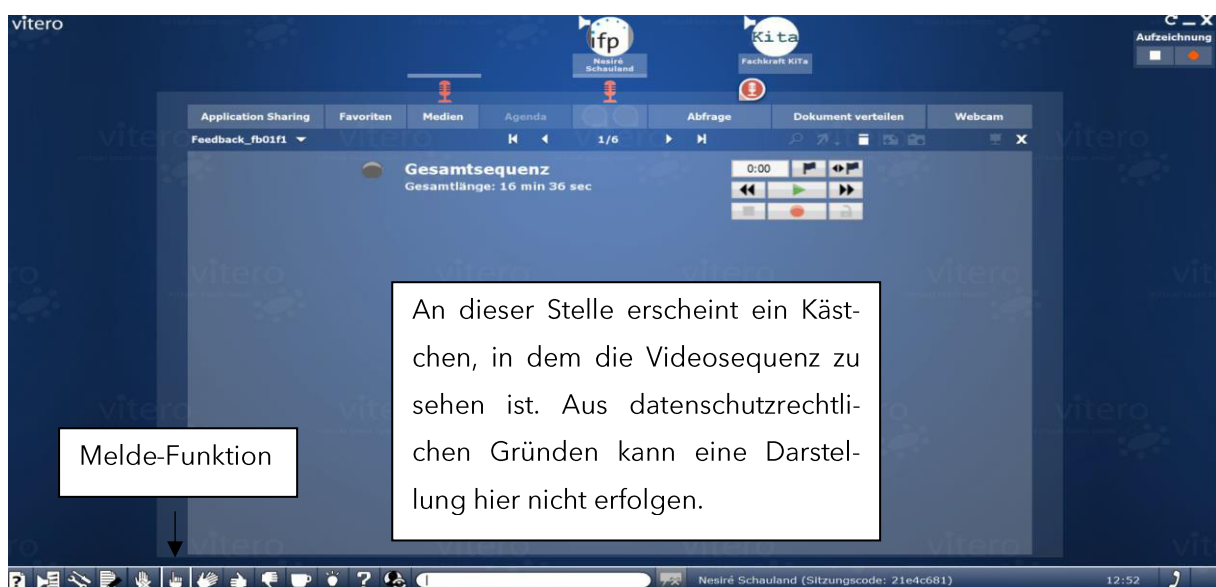


Abbildung 7: Baustein 2 ("How am I going?") und Baustein 3 ("Where to next?")

Am Ende des Feedbackgesprächs wurden mit der Fachkraft aufbauend auf dem, was besprochen wurde, konkrete Ziele in Bezug auf ihr sprachanregendes Verhalten bei der Bilderbuchbetrachtung vereinbart. Über das Menü „Application Sharing“ kann das entsprechende Dokument im „virtual team room“ angezeigt werden, wenn es auf dem Computer des Feedbackgebers geöffnet ist (siehe Abbildung 8). Der Feedbackgeber kann dann die Ziele schriftlich festhalten, die die Fachkraft formuliert hat. Die Fachkraft sieht parallel, was der Feedbackgeber schreibt und kann ggf. Änderungsvorschläge machen.

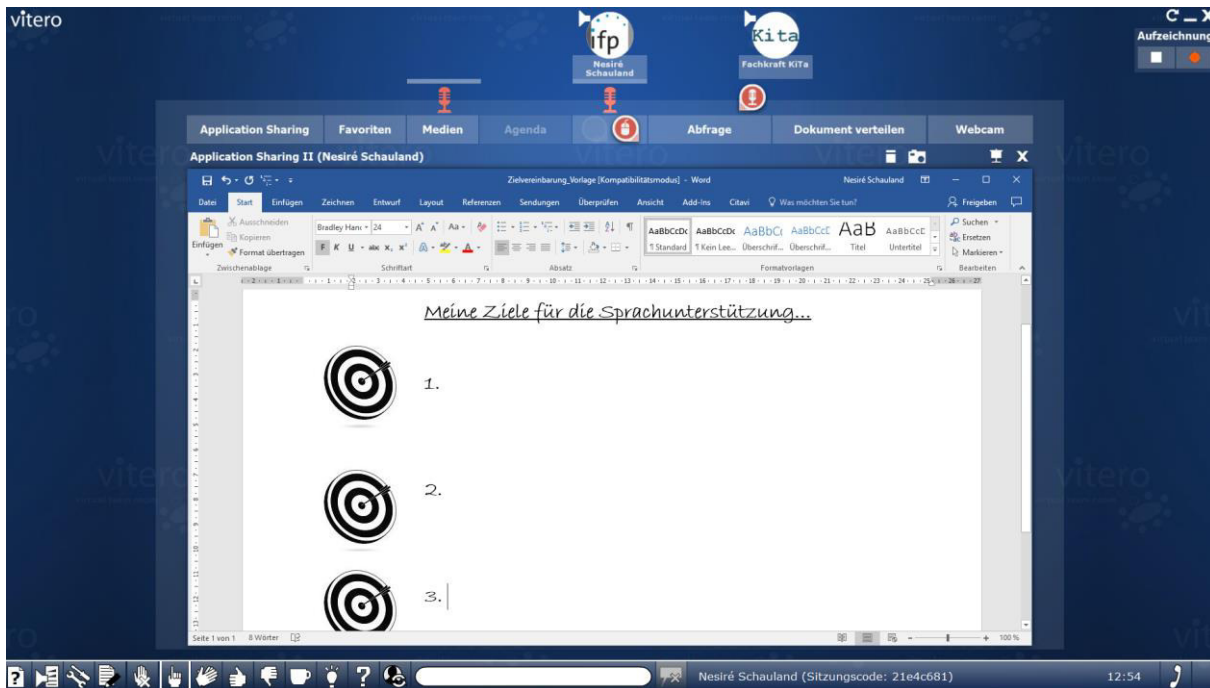


Abbildung 8: Baustein 4: Zielvereinbarung

Das unter dem Fachkraft-Code abgespeicherte Dokument kann schließlich über das Menü „Dokument verteilen“ direkt auf den Computer der Fachkraft übertragen werden (siehe Abbildung 9). So besteht für die Fachkraft zum Beispiel die Möglichkeit, ihre formulierten Ziele direkt auszudrucken.

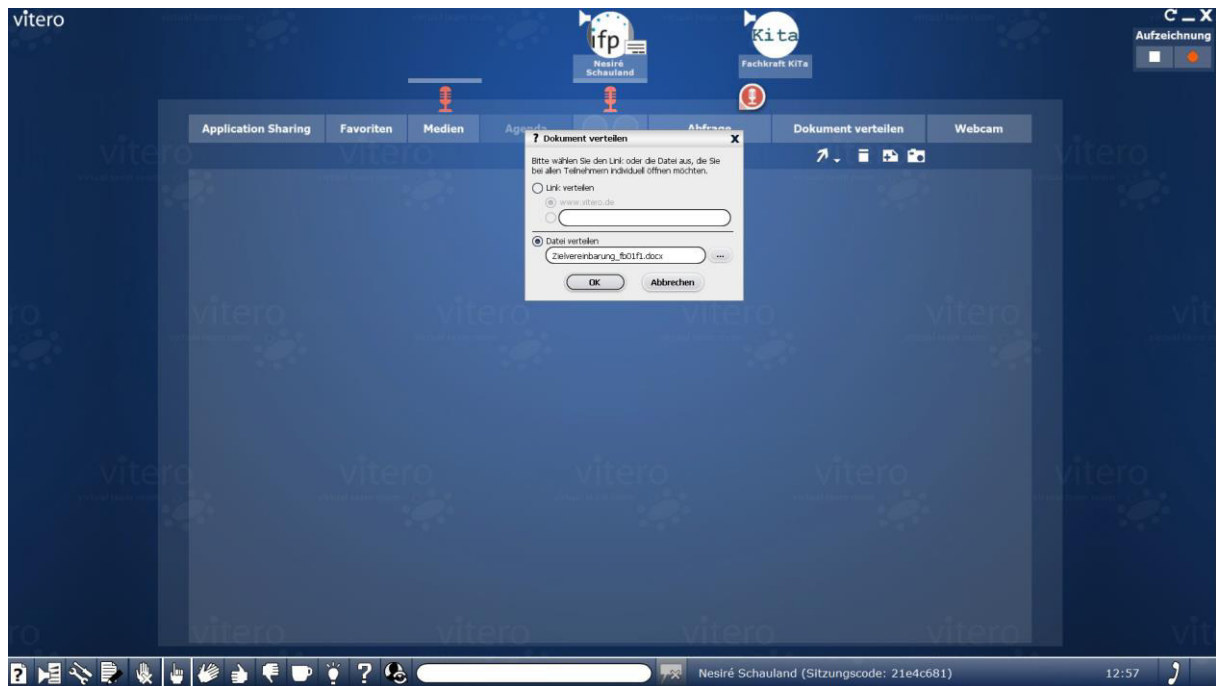


Abbildung 9: Zielvereinbarung an Fachkraft schicken

### 6.3.3 Einfluss- und Kontrollvariablen

In einem schriftlichen Fragebogen zu T1 wurden verschiedene Aspekte abgefragt, die Einfluss auf die Effekte des Feedbacks haben könnten. Welche potenziellen Einflussfaktoren erfasst und wie diese operationalisiert bzw. gemessen wurden, wird im Folgenden vorgestellt.

#### Personenbezogene Merkmale

Im Rahmen des Fragebogens wurden die Fachkräfte gebeten, Angaben zu ihrer Person zu machen. Diese Informationen dienen einerseits der Beschreibung der Stichprobe (siehe Kapitel 6.4), andererseits kann mithilfe dieser Daten sichergestellt werden, dass potenzielle Effekte des Feedbacks auch nach Kontrolle individueller Merkmale bestehen. Tabelle 8 gibt einen Überblick über die verschiedenen Merkmale und wie diese operationalisiert wurden.

Tabelle 8: Erfassung personenbezogener Merkmale in der vorliegenden Studie

Personenbezogene Merkmale	Operationalisierung
Geschlecht	0 = männlich, 1 = weiblich
Alter	Angabe in Jahren
Migrationshintergrund	0 = Nein; 1 = Ja, mind. ein Großelternteil von mir ist im Ausland geboren; 2 = Ja, mind. ein Elternteil von mir ist im Ausland geboren; 3 = Ja, ich bin im Ausland geboren
Mehrsprachigkeit	0 = Nein; 1 = Ja, mit folgenden Sprachen...
Fachlicher/Beruflicher Hintergrund	Berufserfahrung: <ul style="list-style-type: none"> <li>• Anzahl der Jahre, die Fachkraft als pädagogische Fachkraft arbeitet</li> <li>• Anzahl der Jahre, die Fachkraft in der jeweiligen Einrichtung arbeitet</li> </ul> <hr/> Beruflicher Abschluss: <ul style="list-style-type: none"> <li>• Fachschule für Kinderpflege: 0 = Nein, 1 = Ja</li> <li>• Fachschule/ -akademie für Sozialpädagogik / Erzieherin/ Heilpädagogin: 0 = Nein, 1 = Ja</li> <li>• Fachhochschule: 0 = Nein, 1 = Ja; Angabe der Fachrichtung; Abschluss (1 = Bachelor, 2 = Diplom, 3 = Master)</li> <li>• Universität: 0 = Nein, 1 = Ja; Angabe der Fachrichtung; Abschluss (1 = Bachelor, 2 = Diplom, 3 = Master)</li> <li>• Sonstiger Abschluss: 0 = Nein, 1 = Ja; Freie Angabe zur Art dieses Abschlusses</li> </ul> <hr/> Funktion in der Einrichtung: 0 = Nein, 1 = Ja Leitung, stellvertretende Leitung, pädagogische Fachkraft mit Gruppenleitung, pädagogische Fachkraft ohne Gruppenleitung, pädagogische Zweit- oder Ergänzungskraft, Springerkraft, sonstige Funktion (freie Angabe)

### Berufsbezogene Motivation

Die berufsbezogene Motivation der Fachkräfte wurde mit der Kurzversion des „Leistungsmotivationsinventar“ (LMI-K) von Schuler, Prochaska und Frintrup (2001) erhoben. Die Langversion des LMI misst auf den folgenden 17 Dimensionen die Motivation von Personen im beruflichen Kontext: Beharrlichkeit, Dominanz, Engagement, Erfolgsszuversicht, Flexibilität, Flow, Furchtlosigkeit, Internalität, Kompensatorische Anstrengung, Leistungsstolz, Lernbereitschaft, Schwierigkeitspräferenz, Selbständigkeit, Selbstkontrolle, Statusorientierung, Wettbewerbsorientierung, Zielsetzung. Jede der Dimensionen besteht aus zehn Items. Da diese Version eine zu lange Bearbeitungsdauer für die vorliegende Studie gehabt hätte, wurde die Kurzversion mit einem Umfang von 30 Items eingesetzt. Die Items werden auf einer sieben-stufigen Skala eingeschätzt (1 = „trifft gar nicht zu“ bis 7 = „trifft vollständig zu“). Da das Item „Mir sind Aufgaben lieber, die mir leicht von der Hand gehen, ...“ in die



andere Richtung formuliert ist als die restlichen Items, muss es für die weiteren Berechnungen umgepolt werden. Die Werte der 30 Items werden zu einem Summenwert zusammengefasst. Der Summenwert kann anschließend über die Berechnung des Prozentrangs mit der Normierungsstichprobe verglichen werden. In der vorliegenden Studie haben 40.3 % der Fachkräfte einen Prozentrang über 75 und bei ca. zwei Drittel ist der Prozentrang größer als 50. Die berufsbezogene Motivation fällt daher vergleichsweise hoch aus. Die Autoren geben eine interne Konsistenz (Cronbachs Alpha) von 0.94 für die Kurz-Skala an. Mit einem Wert von 0.90 ist das Cronbachs Alpha in der vorliegenden Studie vergleichbar hoch.

### Zielorientierung

Wie in Kapitel 3.2.4 beschrieben wurde, kann in Bezug auf die Ziele von Lernenden zwischen Lern- und Leistungsorientierung unterschieden werden. Wie stark die am Feedback teilnehmenden Fachkräfte lern- bzw. leistungsorientiert sind, wurde mithilfe des „Achievement Goal Questionnaire-Revised“ (AGQ-R) von Elliott und Murayama (2008) gemessen. Der AGQ-R ist eine überarbeitete Fassung des „Achievement Goal Questionnaire“ (AGQ) von Elliot und McGregor (2001). Entsprechend der Differenzierung von Hoska (1993) gehen Elliot und McGregor (2001) davon aus, dass sowohl bei der Lernorientierung („Mastery“) als auch bei der Leistungsorientierung („Performance“) zwischen „acquisition goals“ und „avoidance goals“ differenziert werden muss und entwickelten ein „2 x 2“-Rahmenmodell und den darauf aufbauenden Fragebogen mit den folgenden vier Faktoren: „Mastery-approach goals“, „Mastery-avoidance goals“, „Performance-approach goals“ und „Performance-avoidance goals“. Diese Struktur wurde für den AGQ-R übernommen. Die psychometrischen Eigenschaften des Fragebogens wurden an einer Stichprobe von 229 Studenten geprüft. Die Ergebnisse einer konfirmatorischen Faktorenanalyse bestätigten die angenommene Struktur von vier Faktoren mit jeweils drei Items. Die Items des AGQ-R wurden für die vorliegende Studie von der Autorin ins Deutsche übersetzt und von einer Psychologiestudentin mit Muttersprache Englisch gegenübersetzt. Die Formulierungen wurden an den Kontext der Studie angepasst. Die Fachkräfte wurden entsprechend dem Original gebeten, die Items auf einer Skala von 1 („Stimme gar nicht zu“) bis 5 („Stimme voll und ganz zu“) zu bewerten. In Tabelle 9 sind die in dieser Studie verwendeten Items sowie die Cronbachs Alpha-Werte für die Skalen dargestellt. Zum Vergleich werden ebenfalls die von Elliott und Murayama (2008) angegebenen Cronbachs Alpha-Werte berichtet.

Tabelle 9: Items und Skalen der hier eingesetzten deutschen Übersetzung des AGQ-R sowie die dazugehörigen Cronbachs Alpha-Werte aus der vorliegenden Studie und der Studie von Elliott und Murayama (2008)

Skala	Items	Cronbachs $\alpha$	
		Studie	Autoren
	Es ist mein Ziel, ...		
<b>Mastery-approach goals</b>	...den thematisierten Inhalt vollständig zu beherrschen.	0.56	0.84
	...den erläuterten Inhalt so sorgfältig wie möglich zu verstehen.		
	...so viel wie möglich davon zu lernen.		
<b>Mastery-avoidance goals</b>	...zu vermeiden, weniger zu lernen, als ich könnte.	0.80	0.88
	...ein unvollständiges Verständnis zu vermeiden.		
	...nicht weniger dazu zu lernen als möglich wäre.		
<b>Performance-approach goals</b>	...eine gute Leistung im Verhältnis zu anderen zu zeigen.	0.75	0.92
	...im Vergleich zu anderen gut abzuschneiden.		
	...eine bessere Leistung zu erbringen als die anderen.		
<b>Performance-avoidance goals</b>	...nicht schlechter zu sein als die anderen.	0.93	0.94
	...keine schlechtere Leistung als die anderen zu zeigen.		
	...nicht schlecht abzuschneiden im Vergleich zu anderen.		

Anmerkung zur Stichprobe in der vorliegenden Studie: „Mastery-approach goals“ ( $N = 43$ ), „Mastery-avoidance goals“ ( $N = 40$ ), „Performance-approach goals“ ( $N = 44$ ), „Performance-avoidance goals“ ( $N = 43$ )

Die Skala „Mastery-approach goals“ weist in der vorliegenden Studie keine gute innere Konsistenz auf. Alle anderen Skalen verfügen über ein zufriedenstellendes Cronbachs Alpha. Da es sich beim AGQ-R um ein etabliertes Instrument handelt, wurden die Items dennoch entsprechend dem Original über den Mittelwert zu Variablen zusammengefasst. In Tabelle 10 werden die Mittelwerte und Standardabweichungen der Feedbackgruppe mit den Werten von Elliott und Murayama (2008) verglichen.

Tabelle 10: Deskriptive Kennwerte der Skalen zur Zielorientierung – vergleichend für Feedbackgruppe ( $N = 44$ ) und Vergleichsstichprobe ( $N = 229$ ) von Elliott und Murayama (2008)

Skala	Feedbackgruppe		Vergleichsstichprobe	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Mastery-approach	4.39	0.66	4.23	0.67
Mastery-avoidance	3.88	1.01	3.61	0.95
Performance-approach	2.84	1.07	4.05	0.94
Performance-avoidance	3.17	1.37	3.83	1.06

An den Werten ist zu erkennen, dass die Fachkräfte in der Feedbackgruppe eine ähnlich hohe Lernzielorientierung aufweisen wie die 229 Teilnehmer aus der Vergleichsstichprobe, wobei in beiden Fällen die „Mastery-approach goals“ am stärksten ausgeprägt sind. Im Gegensatz zur Vergleichsstichprobe ist bei den Fachkräften in dieser Studie das Ausmaß der Leistungsorientierung deutlich geringer als das der Lernorientierung.

### Einstellung zu Feedback

Als potenzielle Einflussvariable wurde auch die Einstellung der teilnehmenden Fachkräfte gegenüber Feedback abgefragt. Hierfür wurde die „Feedback Orientation Scale“ (FOS) von Linderbaum und Levy (2010) ins Deutsche übersetzt. In einer Studie mit 280 Studenten (Studie 1) und einer Studie mit 267 Angestellten einer Firma (Studie 2) wurden die psychometrischen Eigenschaften des Fragebogens untersucht. Die Analysen zur faktoriellen Struktur deuten in beiden Untersuchungen auf die folgenden vier Faktoren hin: „Utility“, „Accountability“, „Social Awareness“ und „Feedback Self-Efficacy“. Alle vier Faktoren werden jeweils anhand von fünf Items auf einer fünf-stufigen Skala (1 = „strongly disagree“ bis 5 = „strongly agree“) beschrieben. Für die vorliegende Studie wurden die Items, wie oben beschrieben, übersetzt und an den Kontext angepasst. Tabelle 11 gibt einen Überblick über die deutsche Version der Items und die Cronbachs Alpha-Werte für die entsprechenden Skalen.

Tabelle 11: Items der hier eingesetzten deutschen Übersetzung des FOS sowie die ermittelten Cronbachs Alpha-Werte für alle Skalen - vergleichend für die vorliegende Studie und die Studien von Linderbaum und Levy (2010)

Skala	Items	Cronbachs $\alpha$		
		Studie <sup>1</sup>	Autoren <sup>2</sup>	
Utility	Feedback trägt zu meinem Erfolg in der Arbeit bei.	0.85	0.86	0.88
	Um meine beruflichen Kompetenzen zu entwickeln, bin ich auf Feedback angewiesen.			
	Feedback ist entscheidend, um Leistung zu verbessern.			
	Feedback von Vorgesetzten kann mir dabei helfen, meine Karriere voranzubringen.			
	Ich denke, dass Feedback entscheidend ist, um meine Ziele zu erreichen.			
Accountability	Es ist meine Verantwortung, Feedback umzusetzen, um meine Leistung zu verbessern.	0.86	0.74	0.73
	Es liegt in meiner Verantwortung, angemessen auf Feedback zu reagieren.			
	Es fühlt sich für mich nicht abgeschlossen an, bis ich auf das Feedback reagiere.			
	Wenn mir mein Vorgesetzter Feedback gibt, liegt es in meiner Verantwortung, darauf einzugehen.			
	Ich fühle mich verpflichtet, auf der Grundlage von Feedback Veränderungen vorzunehmen.			
Social Awareness	Ich versuche mir bewusst darüber zu sein, was andere Menschen von mir denken.	0.78	0.80	0.85
	Wenn ich Feedback nutze, bin ich mir bewusster darüber, was Menschen von mir denken.			
	Feedback unterstützt mich darin, zu steuern, welchen Eindruck ich auf andere mache.			
	Feedback lässt mich erkennen, wie ich von anderen wahrgenommen werde.			
	Ich vertraue darauf, dass Feedback mir dabei hilft, einen guten Eindruck zu machen.			
Feedback Self-Efficacy	Ich fühle mich selbstbewusst im Umgang mit Feedback.	0.82	0.77	0.78
	Verglichen mit anderen, bin ich kompetenter im Umgang mit Feedback.			
	Ich glaube, dass ich die Fähigkeit habe, effektiv mit Feedback umzugehen.			
	Ich fühle mich selbstsicher im Umgang mit positivem und negativem Feedback.			
	Ich weiß, dass ich mit dem Feedback umgehen kann, das ich bekomme.			

Anmerkung: <sup>1</sup> „Utility“ (N = 44), „Accountability“, „Social Awareness“ und „Feedback Self-Efficacy“ (N = 42); <sup>2</sup> Werte in der linken Spalte stammen aus Studie 1 und Werte aus der rechten Spalte stammen aus Studie 2

Wie in Tabelle 11 abgelesen werden kann, sind die Cronbachs Alpha-Werte in der vorliegenden Studie vergleichbar mit den Werten der Originalskalen und erreichen eine zufriedenstellende innere Konsistenz. Für die weitere Auswertung wurde jeweils aus allen Items einer Skala der Mittelwert gebildet. Tabelle 12 gibt einen Überblick über die deskriptiven Kennwerte in der Feedbackgruppe und in den Studien von Linderbaum und Levy (2010).

Tabelle 12: Deskriptive Kennwerte der Skalen aus der FOS – vergleichend für Feedbackgruppe ( $N = 44$ ) und Vergleichsstichproben aus Studie 1 ( $N = 280$ ) und Studie 2 ( $N = 267$ )

Skala	Feedbackgruppe		Vergleichsstichproben			
	$M$	$SD$	$M^1$	$SD^1$	$M^2$	$SD^2$
Utility	4.37	0.69	3.92	0.64	3.71	0.84
Accountability	4.35	0.67	3.81	0.54	3.67	0.59
Social Awareness	4.07	0.69	3.73	0.56	3.5	0.75
Feedback Self-Efficacy	3.88	0.77	3.71	0.51	3.74	0.55

Anmerkung: <sup>1</sup> Studie 1; <sup>2</sup> Studie 2

Angesichts der hohen Mittelwerte kann davon ausgegangen werden, dass die am Feedback teilnehmenden Fachkräfte bereits davor eine positive Einstellung gegenüber Feedback hatten. Im Vergleich zeigt die Feedbackgruppe in allen vier Skalen etwas höhere Werte als die Vergleichsstichproben.

### Situationsspezifische Merkmale

Um die Vergleichbarkeit der Situationen (Bilderbuchbetrachtung) von T1 und T2 kontrollieren zu können, wurden im Fragebogen zu T1 und zu T2 verschiedene Angaben zu den Kindern gemacht, die jeweils bei der Bilderbuchbetrachtung anwesend waren. Diese sind in Tabelle 13 dargestellt.

Tabelle 13: Erfassung situationsspezifischer Merkmale in der vorliegenden Studie

Angaben zu den Kindern	Operationalisierung
Anzahl der Kinder und deren Geschlecht	Anzahl von Mädchen und Anzahl von Jungen
Alter der Kinder	Jüngstes Kind: Alter in Jahren und Monaten Ältestes Kind: Alter in Jahren und Monaten
Sprachliche Kompetenz der Kinder	Anzahl der Kinder, die: <ul style="list-style-type: none"> <li>• (noch) nicht sprechen</li> <li>• nicht (Deutsch) sprechen mit Deutsch als Zweitsprache (aber höherer Sprachkompetenz in der Erstsprache)</li> <li>• Sprachanfänger sind (Ein- bis Drei-Wortäußerungen)</li> <li>• Sprachanfänger sind mit Deutsch als Zweitsprache (Ein- bis Drei-Wortäußerungen, höhere Sprachkompetenz in der Erstsprache)</li> <li>• spracherfahren sind (Ausdrucksweise in ganzen Sätzen)</li> <li>• redegewandt sind (differenzierte Ausdrucksweise in komplexen Sätzen, mit gehobenem Wortschatz)</li> </ul>

### Protokoll zum Ablauf des Feedbackgesprächs

In der Arbeitsvorlage, mit der die Feedbackgespräche durchgeführt wurden, haben die Feedbackgeber den Ablauf von jedem Feedback anhand verschiedener Kriterien protokolliert (siehe Kapitel 5.2). In entsprechenden Textfeldern wurden Notizen zu den folgenden Aspekten angefertigt:

- Technische Umsetzung (Gab es technische Probleme?)
- Besprochene Indikatoren
- Vereinbarte Ziele
- Einschätzung des Reflexionsniveaus der Fachkraft.

In Tabelle 14 wird noch einmal zusammenfassend dargestellt, welche Variablen im Rahmen der Studie erfasst wurden. Kontrollgruppe und Feedbackgruppe hatten prinzipiell den gleichen Fragebogen – lediglich die Zielorientierung, die Einstellung gegenüber Feedback und die Bewertung des Feedbackgesprächs wurden in der Kontrollgruppe nicht erfasst.

Tabelle 14: Erfasste Variablen in der vorliegenden Studie

	Feedbackbedingung	Kontrollbedingung
T1	Sprachliche Anregung (Einschätzung von Videoaufnahmen mit CLASS und Kodier-Schema)	
	Personenbezogene Merkmale, Situationsspezifische Merkmale, Berufsbezogene Motivation (Schriftlicher Fragebogen Fachkraft)	
	Zielorientierung (Schriftlicher Fragebogen Fachkraft)	
	Einstellung zu Feedback (Schriftlicher Fragebogen Fachkraft)	
Feed-back-gespräch	Technische Probleme, Reflektierte Indikatoren, Vereinbarte Ziele, Einschätzung des Feedbackgesprächs (Arbeitsvorlage)	
T2	Sprachliche Anregung (Einschätzung von Videoaufnahmen mit CLASS und Kodier-Schema)	
	Situationsspezifische Merkmale (Schriftlicher Fragebogen Fachkraft)	
	Bewertung des Feedbackgesprächs (Schriftlicher Fragebogen Fachkraft)	

## 6.4 Stichprobe

Die Überprüfung der Feedbackeffekte erfolgte an einer Stichprobe von pädagogischen Fachkräften im Kindergarten. In den folgenden Abschnitten wird skizziert, welche notwendige Stichprobengröße zugrunde gelegt und wie die Stichprobe akquiriert wurde. Abschließend wird die für die Untersuchung gewonnene Stichprobe beschrieben.

### Ermittlung der Stichprobengröße

Die notwendige Stichprobengröße für die geplante Untersuchung wurde mittels einer A-Priori-Stichproben-Analyse mit „G-Power“ berechnet. Für die Analyse der Effekte der Feedbackintervention wurde eine Varianzanalyse mit Messwiederholung herangezogen. Da hier vor allem der Interaktionseffekt von Interesse ist, wurde dieser als Grundlage für die Berechnung der Stichprobengröße verwendet. Mehrere Meta-Analysen zu Feedbackeffekten zeigen eine durchschnittliche Effektstärke von  $d = 0.40$  (z.B. Hattie & Timperley, 2007; Klu-

ger & DeNisi, 1996). Dies entspricht einer Effektstärke von  $f = 0.20$ . Ähnliche Effektstärken konnten auch für Weiterbildungen im frühpädagogischen Feld ermittelt werden (Egert, 2015). Die Feedbackstudien wurden hauptsächlich in stark strukturierten Lehr-Lern-Settings mit Wissenstests durchgeführt und bei den Studien zu Weiterbildungseffekten bei Frühpädagogen handelt es sich in der Regel um Untersuchungen, die die Effekte von umfangreichen Weiterbildungen mit mehreren Formaten und Einheiten überprüft haben (z.B. Pianta, Mashburn et al., 2008). Bei den Effekten eines einmaligen Feedbackgesprächs auf das Verhalten der Fachkräfte (Performanz-Ebene) kann daher eher von kleineren Effektstärken ausgegangen werden. Die zu erwartende Effektstärke wurde deshalb auf  $f = 0.15$  festgelegt. Ausgehend von Untersuchungen zur Retest-Reliabilität von CLASS-Beobachtungen (Pianta, La Paro et al., 2008), wird mit einer mittleren Korrelation zwischen den Messungen von  $r = 0.50$  gerechnet. Sollen bei einer Festsetzung der Irrtumswahrscheinlichkeit auf  $\alpha = 0.05$  und einer Teststärke von  $(1-\beta) = 0.80$  kleine bis mittlere Effekte nach Cohen (1988) mit einer Effektstärke von  $f = 0.15$  nachgewiesen werden, wird dazu mindestens eine Stichprobengröße von  $N = 90$  Personen, also 45 Personen pro Bedingung benötigt. Mit einer geplanten Stichprobengröße von 50 Fachkräften pro Gruppe sollten demnach auch (relativ) kleine Effekte nachweisbar sein.

### Stichprobenakquise

Als Grundlage für die Akquise diente eine Liste mit allen Kindertageseinrichtungen in Bayern, die Kinder im Alter zwischen drei und sechs Jahren betreuen. Da die Kitas für die Erhebungen am Erhebungstag gut mit öffentlichen Verkehrsmitteln erreichbar sein sollten, wurden nur Kitas einbezogen, die im Gebiet des Münchner Verkehrs- und Tarifverbundes liegen. Aus diesem Grund waren in der Liste der Kitas insgesamt deutlich mehr Kitas aus dem Stadtgebiet München vertreten als aus umliegenden Landkreisen. Mit dem Ziel, ein möglichst ausgewogenes Verhältnis von städtischen und ländlichen Kitas zu erreichen, fand bei den Ziehungen eine Unterteilung in Kitas aus dem Stadtgebiet München und aus umliegenden Landkreisen statt.

Für die Akquise war es notwendig, die Fachkräfte darüber zu informieren, was im Rahmen der Studie auf sie zukommen würde. Es war daher nicht möglich, erst die Akquise durchzuführen und anschließend eine Randomisierung bezüglich der beiden Bedingungen vorzunehmen. Die Akquise der Fachkräfte erfolgte deshalb getrennt für die Feedback- und die Kontrollbedingung mithilfe einer Zufallsziehung. Die gezogenen Kitas wurden in einem ersten Schritt telefonisch kontaktiert. Dabei wurden die Leitungen der Kitas über das geplante Projekt informiert und um Teilnahme gebeten. Sofern die Leitungen prinzipiell Interesse bekundet haben, wurde eine E-Mail mit einem Überblick über das geplante Vorge-



hen verschickt und ein Termin für ein nächstes Telefongespräch vereinbart. Die Leitungen wurden gebeten, bis zum zweiten Telefonat die Teilnahme am Projekt mit den Mitarbeitern zu besprechen. Als Voraussetzung für die Teilnahme wurde den Kitas genannt, dass die Mitarbeiter/Innen:

- a) als pädagogische Fach- oder Ergänzungskräfte tätig sind (Erzieherin, Berufspraktikanten im Anerkennungsjahr, Kinderpfleger) und
- b) mit Kindern im Alter von drei bis sechs Jahren arbeiten.

Da in der Kontrollgruppe zusätzlich Daten für das Projekt „BiSS-Evaluation“ erfasst wurden, musste in dieser Bedingung auch die Voraussetzung erfüllt sein, dass die teilnehmenden Fachkräfte mindestens drei Vormittage pädagogisch mit den Kindern tätig sind (keine reine Einzelförderung, keine regelmäßigen Abwesenheiten z.B. durch Verwaltungsaufgaben).

Wenn sich pädagogische Fachkräfte bereit erklärt hatten, an der Studie teilzunehmen, wurden telefonisch oder per E-Mail alle wichtigen Informationen abgeklärt.

Da sich die Akquise als äußerst schwierig erwies, wurden in einer zusätzlichen Ziehung gut erreichbare Landkreise außerhalb des MVV-Gebietes einbezogen.

Insgesamt wurden für die Kontrollgruppe 102 Kindertageseinrichtungen telefonisch kontaktiert. Davon konnten 13 Kitas nicht erreicht werden. Von den verbleibenden 89 Kitas haben 67 Kitas die Teilnahme verweigert und 22 Kitas zugesagt. Insgesamt haben sich 61 pädagogische Fachkräfte aus 22 Kitas bereit erklärt, in der Kontrollbedingung an der Studie teilzunehmen. 12 der 61 Fachkräfte haben sich kurzfristig gegen eine Teilnahme entschieden bzw. sind wegen Krankheit ausgefallen. Eine Fachkraft musste aufgrund einer nicht verwertbaren Videoaufnahme von der Auswertung ausgeschlossen werden. In die Auswertung konnten daher Videoaufnahmen von 48 Fachkräften einbezogen werden.

Für die Teilnahme in der Feedbackgruppe wurde zu 150 Kitas telefonisch Kontakt aufgenommen, wobei 16 Leitungen nicht erreicht werden konnten. Von den restlichen 134 Kitas haben 26 Kitas mit insgesamt 53 Fachkräften einer Teilnahme zugestimmt. Von den 53 Fachkräften haben sich 5 Fachkräfte später gegen eine Teilnahme entschieden, sodass ebenfalls 48 Fachkräfte an den Erhebungen teilgenommen haben. Daraus ergibt sich eine Gesamtstichprobe von  $N = 96$  Fachkräfte.

### Beschreibung der Stichprobe

Von den 48 Fachkräften in der Feedbackgruppe fehlen bei vier Fachkräften die Angaben im schriftlichen Fragebogen zu T1. In der Kontrollgruppe haben alle Fachkräfte den Fragebogen ausgefüllt.

In Tabelle 15 werden zentrale demografische Merkmale der Stichprobe getrennt für die beiden Bedingungen vorgestellt. Inwiefern sich die Gruppen hinsichtlich dieser Aspekte statistisch relevant unterscheiden, wird in Kapitel 7.3.3 thematisiert. Der Großteil der teilnehmenden Fachkräfte ist weiblich, hat keinen Migrationshintergrund und ist nicht mehrsprachig, wobei der Anteil an Fachkräften mit Migrationshintergrund und Mehrsprachigkeit in der Feedbackgruppe etwas höher ist. Das Durchschnittsalter liegt bei ca. 39 Jahren.

Tabelle 15: Deskriptive Statistik für Demografische Merkmale der Stichprobe

		Feedbackgruppe	Kontrollgruppe
Geschlecht	Weiblich	46	45
	Männlich	2	3
Alter	<i>M</i>	37.55	39.65
	<i>SD</i>	10.96	12.83
Migrationshintergrund	Nein	30	42
	Ja, mind. ein Großelternteil von mir ist im Ausland geboren	3	1
	Ja, mind. ein Elternteil von mir ist im Ausland geboren	1	1
	Ja, ich bin im Ausland geboren	10	4
Mehrsprachigkeit	Nein	35	43
	Ja	9	5

Anmerkung: Angabe der absoluten Häufigkeit bei Geschlecht, Migrationshintergrund und Mehrsprachigkeit

Der fachliche bzw. berufliche Hintergrund der teilnehmenden Fachkräfte wird in Tabelle 16 anhand der Variablen „Berufserfahrung“, „Beruflicher Abschluss“ und „Funktion in der Einrichtung“ beschrieben. Dabei wird wieder zwischen Kontroll- und Feedbackgruppe unterschieden. In beiden Gruppen haben mehr als die Hälfte der Fachkräfte einen beruflichen Abschluss an einer Fachschule/-akademie für Sozialpädagogik. Die zweithäufigste Qualifikation ist der Abschluss für Kinderpflege. Nur wenige Fachkräfte haben einen höheren Abschluss. Sowohl in der Feedbackgruppe als auch in der Kontrollgruppe arbeiten die Fachkräfte im Durchschnitt ca. fünf Jahre in der Einrichtung. Die Berufserfahrung als pädagogische Fachkraft ist in der Kontrollgruppe mit einem Mittelwert von ca. 14 Jahren höher als in der Feedbackbedingung, in welcher sie zum Zeitpunkt von T1 durchschnittlich 10 Jahre als Fachkraft tätig waren. Hinsichtlich der Funktion in der Einrichtung waren jeweils pädagogische Fachkräfte mit Gruppenleitung am stärksten vertreten.

Tabelle 16: Deskriptive Statistik für den fachlichen/beruflichen Hintergrund der Fachkräfte

			Feedbackgruppe <i>N</i> = 44	Kontrollgruppe <i>N</i> = 48
Beruflicher Abschluss	Fachschule für Kinderpflege	Ja	16	18
		Nein	28	30
	Fachschule/-akademie für Sozialpädagogik etc.	Ja	26	28
		Nein	18	20
	Fachhochschule	Ja	6	4
		Nein	38	44
	Universität	Ja	4	2
		Nein	40	39
	Sonstiger Abschluss	Ja	6	9
		Nein	38	39
Berufserfahrung	Als pädagogische Fachkraft	<i>M</i>	9.86	14.45
		<i>SD</i>	1.50	1.91
	In der Einrichtung	<i>M</i>	4.50	4.69
		<i>SD</i>	0.61	0.79
Funktion in der Einrichtung	Leitung	Ja	3	6
		Nein	41	42
	Stellvertretende Leitung	Ja	9	8
		Nein	35	40
	Päd. Fachkraft mit Gruppenleitung	Ja	17	22
		Nein	27	26
	Päd. Fachkraft ohne Gruppenleitung	Ja	10	11
		Nein	34	37
	Päd. Zweit- oder Ergänzungskraft	Ja	9	14
		Nein	35	34
	Springerkraft	Ja	1	2
		Nein	43	46
	Sonstige Funktion	Ja	6	2
		Nein	38	46

Anmerkung: Angabe der absoluten Häufigkeit beim beruflichen Abschluss und der Funktion in der Einrichtung

Für die Beschreibung der Stichprobe wurden zudem die Leitungen der Kindertageseinrichtungen zum ersten Messzeitpunkt zu Merkmalen und strukturellen Rahmenbedingungen ihrer Einrichtung schriftlich befragt.

Die 96 teilnehmenden Fachkräfte stammen aus insgesamt 45 verschiedenen Einrichtungen, wobei sich die Stichprobe in der Feedbackbedingung auf 24 Kitas und die in der Kontroll-

bedingung auf 21 Kitas verteilt. In Tabelle 17 sind die Angaben zum Träger, zu Einrichtungs- und Betreuungsart und zum organisatorischen Konzept für beide Gruppen zusammengefasst. Manche Fragen wurden nicht von allen Leitungen ausgefüllt, was zu fehlenden Werten in der Darstellung führt. Während in der Feedbackgruppe die Einrichtungen vor allem einen kommunalen oder freien bzw. kirchlichen Träger haben, verteilt sich die Trägerschaft in der Kontrollbedingung relativ gleichmäßig auf die unterschiedlichen Formen. Bei jeweils mehr als der Hälfte der Kitas handelt es sich um altersgemischte Einrichtungen, wobei die Betreuungsart in diesen Einrichtungen variiert. In beiden Gruppen haben 16 Leitungen angegeben, überwiegend in Stammgruppen zu arbeiten und zum Teil eine Öffnung zu haben. Die anderen Konzepte werden nur vereinzelt umgesetzt.

Tabelle 17: Beschreibung der Stichprobe auf Kita-Ebene

		Feedbackgruppe	Kontrollgruppe
Träger	Kommunaler Träger	8	7
	Freier oder kirchlicher Träger	10	4
	Privat-gewerblicher Träger	2	5
	Anderer Träger	3	5
Einrichtungsart	„reine“ Kinderkrippe	2	0
	„reiner“ Kindergarten“	8	5
	Altersgemischte Einrichtung	13	16
Betreuungsart	altersgemischt	9	4
	Teilweise altersgemischt	4	9
	getrennt	6	5
Organisatorisches Konzept	Arbeit in Stammgruppen - ohne Öffnung	2	1
	überwiegend Arbeit in Stammgruppen - zum Teil Öffnung	16	16
	überwiegend offene Arbeit - zum Teil Arbeit in Stammgruppen	3	3
	offene Arbeit - keine Stammgruppen	0	1
	Eingruppige Einrichtung	1	0

Anmerkung: Die Werte stellen die absolute Häufigkeit in der jeweiligen Gruppe dar.

Die Einrichtungsgröße ist in der Kontrollgruppe mit einem Mittelwert von ca. 90 angemeldeten Kindern ( $SD = 33.54$ ) größer als in der Feedbackgruppe, in welcher durchschnittlich 72 Kinder ( $SD = 29.26$ ) betreut werden. In den Kitas der Feedbackgruppe werden von jeder Fachkraft im Durchschnitt 6.5 Kinder ( $SD = 3.02$ ) betreut, in der Kontrollgruppe liegt der Mittelwert bei 6.1 ( $SD = 1.98$ ). Hierbei ist anzumerken, dass sich der Betreuungsschlüssel in den Kitas beider Gruppen stark zwischen den Altersgruppen unterscheidet, wobei er

im U3-Bereich bei ca. 1:3 und bei den Kindern zwischen drei und sechs Jahren bei ca. 1:8 liegt.

## 6.5 Ablauf der Studie

Bevor mit der Stichprobenakquise begonnen werden konnte, wurden alle Träger von Kindertageseinrichtungen in München und Umgebung über das geplante Projekt informiert. Alle freien Träger in Bayern erhielten ein Schreiben mit Informationen zur Studie und der Bitte, Einwände schriftlich zu erheben, sofern vorhanden. Keiner der angeschriebenen Träger formulierte Einwände gegen das Vorhaben. Hinsichtlich der Teilnahme von städtischen Kindertageseinrichtungen der Landeshauptstadt München musste das Forschungsprojekt zunächst beim zuständigen Referat schriftlich beantragt werden. Nachdem der Forschungsantrag genehmigt wurde, konnte mit der Stichprobenakquise begonnen werden.

Sobald sich eine Fachkraft für die Teilnahme an der Studie entschieden hat, wurden per Post alle nötigen Unterlagen an die Kita geschickt. Jede Fachkraft erhielt ein Schreiben, in dem sie noch einmal ausführlich über die Studie informiert wurde. Diesem Schreiben waren zudem eine Information zum Datenschutz sowie eine Einwilligungserklärung beigelegt, welche im Einklang mit dem Bundesdatenschutzgesetz formuliert wurden. Zusätzlich bekam jede Fachkraft eine Kurz-Information, in der der Ablauf der Studie übersichtlich zusammengefasst war. Da auf den Videoaufnahmen nicht nur die Fachkraft, sondern auch die Kinder, die bei der Bilderbuchbetrachtung dabei sind, zu sehen und zu hören sein sollten, war zudem die Einwilligung der Eltern notwendig. Das Paket an die Kita enthielt daher auch Informationsschreiben und Einwilligungserklärungen für die Eltern der Kinder. Die Fachkräfte wurden gebeten, diese an die Eltern zu verteilen und noch vor T1 einzusammeln.

Bei jeder teilnehmenden Fachkraft wurden zwei Videoaufnahmen (T1 und T2) bei einer Bilderbuchbetrachtung angefertigt. Die Termine für die beiden Videoaufnahmen wurden, wenn möglich, so gelegt, dass ca. ein Abstand von vier Wochen dazwischen lag. Um ausreichend Zeit für die Vorbereitung zu haben, fanden die online-basierten Feedbackgespräche in der Feedbackbedingung ca. zwei Wochen nach der ersten Videoaufnahme statt.

Die Datenerhebungen wurden von der Autorin und eigens dafür geschulten Versuchsleitern durchgeführt. Alle Erheber folgten dem in einem Erhebungsleitfaden festgelegten Ablauf (siehe Anhang A). Die Erhebung fand in der jeweiligen Kita statt. Da bei der Kontrollgruppe zusätzlich auch Daten für das Projekt „BiSS-Evaluation“ erfasst wurden, erfolgte in der Kontrollbedingung die Hospitation immer am Vormittag. Den genauen Zeitpunkt für die Bilderbuchbetrachtung konnten alle teilnehmenden Fachkräfte frei wählen.

Die Hospitation zu T1 wurde mit einer kurzen Vorbesprechung begonnen. Diese diente dazu, kurz den Ablauf mit der jeweiligen Fachkraft zu besprechen, die Einwilligung der

Fachkraft sowie der Eltern einzusammeln und gegebenenfalls Fragen zu beantworten. Die Fachkraft wurde gebeten, sicherzustellen, dass bei der videografierten Bilderbuchbetrachtung nur Kinder anwesend sind, von denen eine Einwilligung der Eltern vorliegt. Für die Videoaufnahme erhielt die Fachkraft ein kleines Mikrofon, welches über der Brust an der Kleidung befestigt wurde. Auf diese Weise konnten auch die Äußerungen der Kinder mit einer ausreichend hohen Tonqualität aufgezeichnet werden. Der Empfänger für das Mikrofon befand sich an der Handkamera, welche auf einem Stativ angebracht war. Die Position der Kamera wurde so gewählt, dass möglichst alle Beteiligten (Fachkraft und Kinder) auf der Aufnahme zu sehen sind. Im Anschluss an die Bilderbuchbetrachtung fand eine Nachbesprechung statt. Zum Einstieg wurde die Fachkraft gefragt, wie sie die Hospitation empfunden hat und ob es irgendwelche Besonderheiten an dem Tag gab. Danach wurde die Fachkraft gebeten, den schriftlichen Fragebogen auszufüllen. Abschließend wurden die Details für die zweite Videoaufnahme besprochen. In der Feedbackbedingung wurde zudem, wenn möglich, vom Versuchsleiter die vitero-Software auf dem entsprechenden Computer installiert. In beiden Bedingungen wurde zu T1 der Leitung ein schriftlicher Fragebogen zu strukturellen Rahmenbedingungen übergeben. Nach der ersten Erhebung wurde den Fachkräften aus der Feedbackbedingung eine E-Mail mit allen nötigen Informationen zum Ablauf des Feedbackgesprächs geschickt. Diese enthielt auch die Zugangsdaten für den „virtual team room“ (Sitzungscode und persönliches Passwort) sowie, wenn nötig, eine Anleitung zur Installation von vitero. Eine Webcam und ein Headset wurden entweder bereits zu T1 übergeben oder nach T1 per Post geschickt. Der konkrete Ablauf der Feedbackgespräche ist in Kapitel 5.2 und in Kapitel 6.3.2 ausführlich beschrieben.

Die Hospitation zu T2 erfolgte grundsätzlich nach dem gleichen Schema wie zu T1. Eine Vorbesprechung war nicht mehr zwangsläufig nötig. Die Bilderbuchbetrachtung sollte, wenn möglich, mit dem gleichen oder einem ähnlichen Buch und den gleichen Kindern wie zu T1 durchgeführt werden. In der Nachbesprechung wurden die Fachkräfte gebeten, den T2-Fragebogen auszufüllen. Im Laufe der Hospitation wurde der ausgefüllte schriftliche Fragebogen, den die Leitung zu T1 erhalten hatte, eingesammelt.

Alle teilnehmenden Fachkräfte erhielten nach Abschluss der Studie eine kurze schriftliche Rückmeldung. In dieser wurde vom jeweiligen Beobachter zusammengefasst, was während der Beobachtung positiv aufgefallen ist und wo ggf. noch Optimierungspotenzial gesehen wird.

## 6.6 Statistische Auswertung

Ein Überblick über das Vorgehen zur Überprüfung der in Kapitel 6.1 formulierten Fragestellungen und Hypothesen soll in den folgenden Abschnitten gegeben werden. Alle

Auswertungen wurden mit dem Statistikprogramm IBM SPSS Statistics Version 24 oder manuell durchgeführt. Für die gesamten Berechnungen wird entsprechend der gebräuchlichen Vorgehensweise das Signifikanzniveau so interpretiert, dass  $p < 0.01$  ein hoch signifikantes,  $p < 0.05$  ein signifikantes und  $p > 0.05$  ein nicht signifikantes (n.s.) Ergebnis darstellt.

### 6.6.1 Beurteilerübereinstimmung und -reliabilität

Um die Effekte des Feedbacks auf das sprachliche Anregungsverhalten der Fachkräfte untersuchen zu können, wurden die zu T1 und T2 angefertigten Videoaufnahmen von vier Beurteilern mit der CLASS und dem entwickelten Kodier-Schema ausgewertet. Die Übereinstimmung bzw. Reliabilität zwischen den Ratern wurde anhand einer Teilstichprobe der Videoaufnahmen überprüft. Dafür wurden von jedem Rater jeweils drei Videos von den anderen drei Ratern eingeschätzt. Es wurden demnach von jedem Rater neun Videos doppelt kodiert. Insgesamt ergibt dies eine Anzahl von 36 doppelt kodierten Videoaufnahmen, was 18,75 % der Videoaufnahmen insgesamt darstellt.

In Anlehnung an die Vorgaben des CLASS-Manuals (Pianta, La Paro et al., 2008) wurde für die Bewertung der Dimension „Language Modeling“ eine Differenz von  $\pm 1$  auf der siebenstufigen Skala als Übereinstimmung festgelegt. Als Maß für die Beurteilerübereinstimmung gilt hier die prozentuale Übereinstimmung (PÜ) zwischen den Ratern (Wirtz & Caspar, 2002), deren Berechnung in der folgenden Gleichung dargestellt ist.

Formel 1: Berechnung der prozentualen Beurteilerübereinstimmung

$$P\ddot{U} = \frac{\text{Anzahl der beobachteten Übereinstimmungen}}{N} \times 100 \%$$

Da es sich bei der mithilfe des Kodier-Schemas ermittelten Anzahl der sprachlichen Verhaltensweisen um intervallskalierte Werte handelt, eignet sich in diesem Fall die Intraklassenkorrelation (ICC) am besten für die Bestimmung der Interraterreliabilität (Wirtz & Caspar, 2002). Die Intraklassenkorrelation beruht auf den Annahmen des varianzanalytischen Modells und gibt daher Auskunft über den Anteil der Varianz in den Beurteilungen unterschiedlicher Rater, der durch die wahren Merkmalsausprägungen aufgeklärt werden kann (Shrout & Fleiss, 1979). Der Intraklassenkorrelations-Koeffizient kann Werte zwischen -1 und 1 annehmen und kann ähnlich interpretiert werden wie die Produkt-Moment-Korrelation. Je mehr sich der Koeffizient dem Wert 1 nähert, desto höher ist der Zusammenhang zwischen den Urteilen der Rater und desto höher ist die Reliabilität der Einschätzungen. Da der ICC-Koeffizient ein Maß für die Varianzaufklärung ist, werden negative Werte nicht als Reliabili-

tätsmaß angegeben, sondern als eine Reliabilität von 0 (Varianzaufklärung von 0 %) interpretiert (Wirtz & Caspar, 2002). Eine Intraklassenkorrelation von 0.70 wird in der Regel als ein Hinweis auf eine gute Reliabilität betrachtet. Wirtz und Caspar (2002) merken jedoch diesbezüglich an, dass die Intraklassenkorrelation immer in Abhängigkeit der gemessenen Variable und der untersuchten Stichprobe interpretiert werden muss. So hängt die Höhe der Reliabilität zum Beispiel von den vorhandenen Mittelwertunterschieden zwischen den Personen ab. Wenn nur wenig Varianz in den wahren Werten vorhanden ist, ist es deutlich schwieriger, diese Varianz aufzuklären.

Für die Berechnung des ICC-Koeffizienten gibt es verschiedene Möglichkeiten. Bei der Auswahl der geeigneten Variante werden unterschiedliche Kriterien herangezogen. Da in der vorliegenden Arbeit nicht alle 36 Videoaufnahmen von den gleichen Ratern eingeschätzt wurden, muss ein einfaktorielles Modell verwendet werden (Shrout & Fleiss, 1979). Beim einfaktoriellen Modell wird automatisch eine unjustierte Intraklassenkorrelation berechnet. Der unjustierte ICC-Koeffizient berücksichtigt Mittelwertunterschiede zwischen den Ratern für die Reliabilitätsschätzung. Für eine hohe Reliabilität müssen demnach die absoluten Werte, die durch die Rater vergeben wurden, ähnlich sein. Im Vergleich dazu muss bei einer justierten ICC hingegen nur die relative Position der geschätzten Werte zum Mittelwert der Rater vergleichbar sein (Wirtz & Caspar, 2002). Da für die vorliegende Studie die durchschnittliche Reliabilität der Einschätzungen durch das Kodier-Schema von Interesse ist, wird der ICC-Koeffizient für den Mittelwert über die Rater herangezogen.

Im hier verwendeten einfaktoriellen Modell kann der Wert der Intraklassenkorrelation bei nicht homogenen Varianzen zwischen den Ratern verzerrt werden (Wirtz & Caspar, 2002). Aus diesem Grund wurde vorher die Varianzhomogenität für alle Variablen des Kodier-Schemas getestet. Die Homogenität von Varianzen wird häufig mit dem sogenannten Levene-Test überprüft. Wird das zuvor festgelegte Signifikanzniveau im Levene Test auf Fehlervarianzen verfehlt, so kann man von Varianzhomogenität ausgehen. Der Levene-Test wird allerdings in der Literatur für seine Anfälligkeit, vor allem gegenüber kleinen Stichproben oder schief verteilten Variablen, kritisiert. Aus diesem Grund wurde die Varianzhomogenität mithilfe des  $F$ -max-Tests von Hartley (1950) untersucht, welcher die größten Varianz mit der kleinsten Varianz vergleicht. Das Verhältnis der beiden Varianzen liegt im Idealfall nah an 1. Der errechnete Wert wird mit dem  $F_{max}$ -Wert verglichen, ab dem das Verhältnis der beiden Varianzen überzufällig von 1 abweicht (bei einer angenommenen Wahrscheinlichkeit des Fehlers 1. Art von 0.05), also keine Varianzhomogenität vorliegt. Ist der errechnete  $F$ -Wert kleiner als das kritische Varianzverhältnis, kann von Varianzhomogenität ausgegangen werden. Das kritische Varianzverhältnis ist von der Anzahl der Gruppen ( $k$ ) und der Anzahl der Beobachtungen pro Gruppe bzw. den daraus ermittelten Freiheitsgraden ( $df = N-1$ ) ab-



hängig. Da an dieser Stelle die Varianzhomogenität zwischen den Ratern untersucht werden soll, stellt jeder Rater eine Gruppe mit jeweils der gleichen Anzahl an Beobachtungen dar. Die Varianzhomogenität wurde immer zwischen zwei Ratern für die sechs doppelt kodierten Videoaufnahmen geprüft. Bei insgesamt vier Ratern ergibt dies sechs Rater-Paare, für die jeweils für alle Variablen des Kodier-Schemas ein  $F$ -Wert aus dem Verhältnis der beiden Varianzen berechnet wurde. Der kritische  $F$ -Wert für zwei Gruppen (Rater) sowie sechs Beobachtungen und somit fünf Freiheitsgraden liegt laut Hartley (1950) bei 7.15, das heißt die größere der beiden Varianzen dürfte für die Annahme der Varianzhomogenität maximal 7.15-fach größer sein als die kleinere Varianz. Die diesbezügliche Auswertung ergab, dass bei allen sechs Rater-Paaren das Varianzverhältnis bei dem größten Teil der Variablen den kritischen Wert unterschritten hat. Da es sich nur um einzelne Abweichungen handelt, die zudem nicht systematisch bei bestimmten Variablen auftreten, konnte die Varianzhomogenität bestätigt und somit der ICC-Koeffizient für alle Variablen berechnet werden. Die Ergebnisse werden in Kapitel 1417.3.2 beschrieben.

## 6.6.2 Variablenbildung

Über das sprachliche Anregungsverhalten hinaus wurden zur Überprüfung der Fragestellungen und Hypothesen verschiedene andere Instrumente und Items eingesetzt (siehe Kapitel 6.3). Zum Teil war es notwendig, aus den erfassten Informationen neue Variablen bzw. Skalen zu bilden. Skalen aus bestehenden Instrumenten bzw. Items oder Variablen, die entsprechend ihrer Operationalisierung im Fragebogen verwendet werden konnten, werden an dieser Stelle nicht noch einmal erwähnt.

### Bewertung des Feedbackgesprächs durch die Feedbackgeber

Aus den Notizen in der Arbeitsvorlage (siehe Kapitel 5.2.3) zu den technischen Problemen beim Feedbackgespräch und dem wahrgenommenen Reflexionsniveau der Fachkraft wurden neue Variablen gebildet.

Hinsichtlich der technischen Probleme wurde zunächst in einer Variable („Techn\_Prob“) kodiert, ob überhaupt technische Schwierigkeiten aufgetreten sind (0 = „nein“, 1 = „ja“). In einem zweiten Schritt wurden die genannten Probleme kategorisiert. Folgende Kategorien sind dabei entstanden:

- Schlechte Ton-Übertragung
- Kein Ton
- Keine Webcam-Übertragung
- Probleme bei Übertragung der Videosequenzen
- Unterbrechung der Verbindung/Absturz von vitero
- Feedbackgespräch nicht in Kita möglich

Für jedes Feedbackgespräch wurde schließlich angegeben, welche dieser technischen Probleme aufgetreten sind. Da es zum Teil Mehrfachnennungen gab, wurde für jede der oben genannten Kategorien eine Variable gebildet. Bei jedem Feedbackgespräch mit technischen Schwierigkeiten wurde für alle Kategorien bewertet, ob das entsprechende Problem vorkam oder nicht (0 = „nein“, 1 = „ja“).

Das Reflexionsniveau der Fachkraft wurde vom Feedbackgeber einer von drei Stufen zugeordnet: 1 = „niedrig“, 2 = „mittel“ und 3 = „hoch“. Ein Beispiel für eine niedrige Ausprägung ist es, wenn die Fachkraft während des Feedbackgesprächs eine Abwehrhaltung zeigte und kein Optimierungspotenzial erkannt hat. Fachkräfte im mittleren Niveau waren grundsätzlich in der Lage ihr Verhalten zu reflektieren, hatten aber unter anderem Schwierigkeiten konkrete Beispiele zu erkennen oder Lösungen zu finden. Fachkräfte auf der höchsten Stufe konnten weitestgehend eigenständig Beispiele mit Optimierungspotenzial identifizieren und Lösungsmöglichkeiten erarbeiten.

### Störvariablen

Der berufliche Abschluss sowie die Funktion in der Einrichtung wurden für weitere Berechnungen so umkodiert, dass eine neue Variable aus den bestehenden Mehrfachantworten gebildet wurde, wobei jeweils der höchste Abschluss bzw. die höchste Funktion einer Fachkraft gezählt wurde.

Obwohl es in der Versuchsplanung vorgegeben war, dass die Erhebungen im Abstand von vier Wochen stattfinden und die Situationen von T1 und T2 vergleichbar sein sollen (Anzahl, Alter und Sprachniveau der Kinder bei der Bilderbuchbetrachtung), war dies aufgrund organisatorischer Gründe für die Kitas nicht immer möglich. Um eventuelle Abweichungen bei der Auswertung berücksichtigen zu können, wurden für die entsprechenden Variablen neue Variablen mit den jeweiligen Differenzwerten zwischen T1 und T2 berechnet.

### **6.6.3 Überprüfung der Hypothesen und Fragestellungen**

Die Überprüfung der Hypothesen und Fragestellungen erfolgte mithilfe verschiedener inferenzstatistischer Verfahren. Da es sich hierbei um gängige Auswertungsmethoden handelt,

werden im Folgenden die eingesetzten Signifikanztests nur zusammenfassend skizziert. Eine detaillierte Beschreibung der Testverfahren findet sich beispielsweise bei Bühner und Ziegler (2017).

### Effekte des Feedbacks auf das sprachliche Interaktionsverhalten

Mithilfe von Varianzanalysen nach dem Allgemeinen Linearen Modell können Mittelwertunterschiede intervallskalierter Variablen bei mehr als zwei Gruppen auf Signifikanz getestet werden. Zur Untersuchung der Effekte des Feedbackgesprächs wurden für die verschiedenen Aspekte der sprachlichen Anregung zweifaktorielle Varianzanalysen mit Messwiederholung (gemischtes Design) berechnet. Dabei wurde die Untersuchungsbedingung (Feedback- oder Kontrollbedingung) als Gruppenfaktor in das Modell aufgenommen. Die zwei Messungen (T1 und T2) stellen den zweiten Faktor dar (Messwiederholung). In der Varianzanalyse wird geprüft, in welchem Ausmaß die Gesamtvarianz der abhängigen Variable (sprachliche Anregung) durch den Einfluss der unterschiedlichen Ausprägungen der unabhängigen Variable erklärt werden kann. Bei der Varianzanalyse mit Messwiederholung werden zwei Haupteffekte und ein Interaktionseffekt getestet. Der erste Haupteffekt überprüft, ob es Unterschiede in der sprachlichen Anregung zwischen der Feedback- und der Kontrollgruppe gibt (unabhängig vom Messzeitpunkt). Mithilfe des zweiten Haupteffektes wird untersucht, ob sich die sprachliche Anregung von T1 zu T2 verändert hat (unabhängig von der Gruppe). Der Interaktionseffekt ist das eigentlich wesentliche Ergebnis für die Überprüfung der Feedback-Effekte. Er gibt Auskunft darüber, inwiefern die Veränderung der sprachlichen Anregung von der Teilnahme am Feedbackgespräch abhängig ist. Nach Bühner und Ziegler (2017) gelten für die zweifaktorielle Varianzanalyse mit Messwiederholung vor allem folgende Voraussetzungen:

- Normalverteilung der kontinuierlichen Variablen in allen Teilstichproben
- Homogenität der Gruppenvarianzen
- Homogenität der Varianzen und Kovarianzen der Messwiederholungen (Sphärizität).

Bezüglich der Varianzhomogenität für den Messwiederholungsfaktor wird in der Regel der Mauchly-Test auf Sphärizität herangezogen. Unter Sphärizität versteht man, dass die Varianzen der Differenzen zwischen den Messzeitpunkten gleich hoch sind (Bühner & Ziegler, 2017). Da in der vorliegenden Studie der Innersubjektfaktor „Messzeitpunkt“ nur zwei Stufen hat, gibt es nur eine Differenz und Sphärizität ist daher automatisch gegeben.

Um die Homogenität der Varianzen zwischen den Gruppen zu testen, wurde erneut der  $F$ -max-Test durchgeführt. Dafür wurde die Varianz der Variablen zur sprachlichen Anregung in der Feedbackgruppe mit der Varianz der Werte in der Kontrollgruppe verglichen. Als

Datengrundlage dienten die Einschätzungen zu T1. Da es sich in diesem Fall um zwei Gruppen mit jeweils 48 Personen pro Gruppe ( $df = 47$ ) handelt, hat die Prüfgröße  $F_{\max}$  einen Wert von 2.08. Bei den Variablen „Längere Gespräche“, „Erzählaufforderungen ohne Antwortmöglichkeit“, „Selftalk“ sowie „Erweiterung mit Korrektiv“ wurde das kritische Varianzverhältnis überschritten. In allen übrigen Variablen kann von Varianzhomogenität ausgegangen werden.

Aussagen über die Normalverteilung der abhängigen Variablen können über die Interpretation der Histogramme gemacht werden. Die Sichtung der Histogramme für die Verteilung der Werte zur sprachlichen Anregung zeigte für beide Gruppen zu T1 und T2, dass die Variable „Language Modeling“ nicht stark von der Normalverteilung abweicht. Die Werte für die einzelnen sprachanregenden Verhaltensweisen, welche mit dem Kodier-Schema ermittelt wurden, weisen durchweg eine linkssteile und rechtsschiefe Verteilung auf.

Bühner und Ziegler (2017) merken an, dass die Varianzanalyse relativ robust gegenüber Verletzungen der Normalverteilungsannahme und Varianzhomogenität ist, wenn die Gruppen nicht ungleich groß sind. Da in der vorliegenden Studie in beiden Bedingungen jeweils 48 Fachkräfte teilgenommen haben und bei jeder Fachkraft alle Daten zu T1 und T2 vorliegen, wird die Varianzanalyse, trotz zum Teil fehlender Normalverteilung und Varianzhomogenität, angewendet.

Als Maß für die Größe der Effekte wird das partielle  $\eta^2$  ( $\eta_p^2$ ) verwendet (Bühner & Ziegler, 2017). Die Effektstärke wird nach den Konventionen von Cohen (1988) interpretiert (kleiner Effekt:  $\eta_p^2 = 0.01$ , mittlerer Effekt:  $\eta_p^2 = 0.06$  und großer Effekt:  $\eta_p^2 = 0.14$ ). Wird  $\eta_p^2$  mit 100 multipliziert, ergibt sich der Prozentanteil der Varianz in der abhängigen Variable, welcher durch den jeweiligen Effekt erklärt werden kann.

Der Interaktionseffekt der Varianzanalyse sagt aus, ob sich die Fachkräfte in der Feedbackbedingung zwischen T1 und T2 signifikant stärker in ihrem Verhalten verändert haben als die Fachkräfte in der Kontrollbedingung. Um herauszufinden, wie groß die beobachteten Verhaltensunterschiede jeweils in den beiden Gruppen sind, wurden  $t$ -Tests für abhängige Stichproben getrennt für die Feedback- und die Kontrollbedingung berechnet. Für die Durchführung von  $t$ -Tests gilt ebenfalls die Normalverteilung als Voraussetzung. Beim  $t$ -Test für abhängige Stichproben bezieht sich die Normalverteilung jedoch nicht auf die Messwerte, sondern auf die Differenzen der Messwerte zwischen den beiden Messzeitpunkten. Zur Überprüfung der Normalverteilung wurde daher aus den Werten der Variable „Language Modeling“ und den der sprachanregenden Verhaltensweisen jeweils eine neue Variable für die Messwertdifferenz berechnet. Da bei keiner der Differenzvariablen starke Abweichungen zur Normalverteilung vorliegen, konnte der  $t$ -Test für abhängige Stichproben eingesetzt werden. Zur Interpretation der Größe der Mittelwertdifferenz im  $t$ -Test für abhängi-

ge Stichproben wurde Hedges  $g$  bzw. Cohens  $d_z$  manuell berechnet. Die Formel dafür lautet wie folgt.

Formel 2: Berechnung der Effektstärke  $g$  bzw.  $d_z$

$$g = d_z = \frac{\hat{\mu}_1 - \hat{\mu}_2}{\hat{\sigma}_D}$$

$d_z$  = Effektstärke für abhängige Stichproben von Cohen

$\hat{\mu}_1$  = Mittelwert zu T1

$\hat{\mu}_2$  = Mittelwert zu T2

$\hat{\sigma}_D$  = Standardabweichung der Differenz

Cohen (1988) klassifiziert die Effektstärke  $d_z$  ( $g$ ) in folgenden Stufen:

- Kleiner Effekt:  $d_z = 0.20$
- Mittlerer Effekt:  $d_z = 0.50$
- Großer Effekt:  $d_z = 0.80$ .

Mithilfe der zweifaktoriellen Varianzanalysen mit Messwiederholung wurden die Mittelwertunterschiede zwischen T1 und T2 in Abhängigkeit von der Teilnahme am Feedback untersucht. Diese Vorgehensweise gibt Aufschluss über die allgemeine Effektivität des Feedbacks. Es können jedoch daraus keine Aussagen über die individuelle Wirksamkeit abgeleitet werden. Um herauszufinden, wie viele Fachkräfte tatsächlich signifikant vom Feedback profitieren, kann der sogenannte „Reliable Change Index“ (RC) nach Jacobson und Truax (1991) berechnet werden. Die dafür eingesetzten Formeln werden im Folgenden dargestellt.

Formel 3: Berechnung des "Reliable Change Index" nach Jacobson und Truax (1991)

$$RC = \frac{x_2 - x_1}{S_{diff}}$$

$$S_{diff} = \sqrt{2(S_{err})^2}$$

$$S_{err} = s_1 \sqrt{1 - r}$$

$x_1$  = Wert vor Feedback

$x_2$  = Wert nach Feedback

$S_{diff}$  = Standardfehler der Differenz

$S_{err}$  = Standardmessfehler

$s_1$  = Standardabweichung zu T1

$r$  = Reliabilität der Variable

Die für den Standardmessfehler verwendete Standardabweichung bezieht sich auf die Werte der gesamten Stichprobe zu T1 (Kontroll- und Feedbackbedingung). Für die Reliabilität der CLASS wurde der Wert aus dem CLASS-Manual für den Bereich „Instructional Support“ (0.86) herangezogen, da für die Dimension „Language Modeling“ kein Wert angegeben ist. In Bezug auf die Anzahl der sprachlichen Verhaltensweisen wird auf die ICC-Werte der jeweiligen Variable aus der vorliegenden Studie zurückgegriffen (siehe Kapitel 7.3.2). Der RC-Wert wird als z-Wert interpretiert, das heißt ein  $RC > 1.96$  spricht für eine statistisch bedeutende Veränderung ( $p < 0.05$ ). Mithilfe der beschriebenen Formeln wurde für alle Fachkräfte der Feedbackgruppe der „Reliable Change Index“ für jede Variable der sprachlichen Anregung berechnet. Anschließend wurde die Anzahl der Fachkräfte ermittelt, bei denen der RC-Wert  $> 1.96$  ist und somit von einer reliablen Veränderung ausgegangen werden kann. Multipliziert man den Standardfehler mit 1.96, erhält man den Differenzwert, der aufgrund des Zufalls zu erwarten wäre. Ist der Differenzwert einer Person größer als dieser, deutet dies auf eine relevante Veränderung hin. Da die Daten in der vorliegenden Studie zum Teil eine deutlich größere Streuung aufweisen als in anderen Studien (siehe Kapitel 7.3.1) und der „Reliable Change Index“ stark von der angenommenen Standardabweichung abhängt, wurde dieser zusätzlich mit einem kleineren Konfidenzintervall interpretiert. Ausgehend von einer Sicherheitswahrscheinlichkeit von 10 % gelten alle RC-Werte  $> 1.64$  als tendenziell reliable Veränderung.

### Einfluss von anderen Variablen auf die Effekte des Feedbacks

Da Unterschiede zwischen der Kontroll- und der Feedbackgruppe einen Einfluss auf die Effekte des Feedbacks haben könnten, wurde zunächst geprüft, ob bedeutende Unterschiede vorliegen und anschließend die Wirkung des Feedbacks, sofern signifikante Effekte festgestellt wurden, unter Kontrolle bzw. Einfluss dieser Variablen noch mal untersucht.

In Kapitel 6.4 wurde die Stichprobe getrennt für beide Gruppen anhand verschiedener personenbezogener Merkmale beschrieben. Auch unterschiedliche situative Bedingungen könnten die Effekte verzerren. Um herauszufinden, ob sich die Gruppen in diesen Faktoren statistisch relevant voneinander unterscheiden, wurden Chi-Quadrat-Tests bei kategorialen Variablen (Geschlecht, Migrationshintergrund, Mehrsprachigkeit, Beruflicher Abschluss, Funktion in der Einrichtung) sowie  $t$ -Tests für unabhängige Stichproben bei intervallskalierten Variablen (Alter, Berufserfahrung, Motivation, situative Unterschiede zwischen T1 und T2, Zeitabstand zwischen den Erhebungen, Leistung zu T1) berechnet. Zusätzlich zur Normalverteilung wird beim  $t$ -Test für unabhängige Stichproben die Varianzhomogenität zwischen den Gruppen vorausgesetzt. Die Homogenität der Varianzen in Bezug auf die potenziellen Störvariablen wurde, wie oben beschrieben, mit dem  $F$ -max-Test überprüft. Bühner und Ziegler (2017) weisen darauf hin, dass der  $t$ -Test trotz Verletzungen der Voraussetzungen den alternativen nonparametrischen Testverfahren vorgezogen werden sollte. Nur in Ausnahmefällen (Varianzen unterscheiden sich stark voneinander, keine Normalverteilung, Gruppen sind sehr ungleich groß,  $p$ -Wert liegt knapp unter der Signifikanzgrenze) sollte nach Ansicht der Autoren der Mann-Whitney-U-Test als nonparametrische Alternative gewählt werden. Da dies nicht der Fall ist, konnte der  $t$ -Test für unabhängige Stichproben verwendet werden. Im Falle inhomogener Varianzen wird auf die Kennwerte des Welch-Tests mit der entsprechenden Korrektur der Freiheitsgrade zurückgegriffen. Da an dieser Stelle die Nullhypothese die Wunschhypothese darstellt (es gibt keinen Unterschied zwischen Feedback- und Kontrollbedingung in den potenziellen Störvariablen), muss der Fehler 2. Art kontrolliert werden (Beta-Fehler). Das Signifikanzniveau wird daher sowohl bei den Chi-Quadrat-Tests als auch bei den  $t$ -Tests für unabhängige Stichproben auf 20 % festgesetzt. Beim Chi-Quadrat-Test wird die exakte Alpha-Fehler Wahrscheinlichkeit als Signifikanzwert verwendet. Für die Unterschiede in kategorialen Variablen geben die Koeffizienten Cramers  $V$  oder Phi ( $\phi$ ) Aufschluss über die Stärke des Effekts. In beiden Fällen handelt es sich um ein Zusammenhangsmaß für kategoriale Variablen und kann daher wie  $r$  interpretiert werden. Nach Cohen (1988) gelten dabei folgende Konventionen: 0.10 = kleiner Effekt, 0.30 = mittlerer Effekt und 0.50 = großer Effekt.

Als Maß für die Effektstärke in Bezug auf Unterschiede zwischen unabhängigen Stichproben wird Hedges  $g$  verwendet. Anstatt der geschätzten Standardabweichung der Differenz wird in diesem Fall allerdings die geschätzte Standardabweichung für die Grundgesamtheit verwendet. Bei unterschiedlich großen Gruppen muss dafür die gepoolte Standardabweichung ( $\hat{\sigma}_p$ ) berechnet werden, welche sich aus der folgenden Formel ergibt.

Formel 4: Berechnung der gepoolten Standardabweichung

$$\hat{\sigma}_p = \sqrt{\frac{(n_1 - 1) \times \hat{\sigma}_1^2 + (n_2 - 1) \times \hat{\sigma}_2^2}{n_1 + n_2 - 2}}$$

Handelt es sich um gleich große Gruppen, kann demnach der Mittelwert aus beiden Standardabweichungen herangezogen werden. Zur Ermittlung von  $g$  wird schließlich die Differenz der Mittelwerte durch die gepoolte Standardabweichung geteilt. Die Interpretation erfolgt wie beim  $t$ -Test für abhängige Stichproben.

Zur Kontrolle von intervallskalierten Störvariablen können Kovarianzanalysen berechnet werden. Um die Effekte des Feedbacks unter Einfluss von kategorialen Störvariablen zu analysieren, wurde die kategoriale Störvariable als zusätzlicher Faktor mit in das Modell der zweifaktoriellen Varianzanalyse mit Messwiederholung aufgenommen. Die Voraussetzungen wurden bereits beschrieben. Der Interaktionseffekt zwischen Messwiederholung, Bedingung und Störvariable gibt in diesem Fall Auskunft über die Wirkung der Störvariablen auf die Effekte des Feedbacks.

Wie in Kapitel 3.2.4 erläutert, können unterschiedliche Faktoren den Erfolg von Feedback beeinflussen. Angelehnt daran wurden in Abschnitt 6.1 Hypothesen und Fragestellungen zum Einfluss verschiedener Faktoren (Zielorientierung, Motivation, Einstellung gegenüber Feedback, Reflexionsniveau der Fachkräfte, Bewertung des Feedbacks durch Fachkräfte, technische Probleme) formuliert. Zur Überprüfung dieser wurden hierarchische Regressionsanalysen durchgeführt. Dabei wurden die in Kapitel 6.3.1 und 6.3.3 beschriebenen Variablen als Prädiktoren zur Vorhersage der Veränderung zwischen T1 und T2 (Kriterium) eingesetzt. Als Kriterium wurden nur Variablen einbezogen, bei denen signifikante Effekte des Feedbacks gefunden wurden. Für jede dieser Variablen wurde die Differenz zwischen T1 und T2 ( $T2 - T1$ ) berechnet, wobei positive Werte für eine Verbesserung stehen. Da untersucht werden soll, ob die teilnehmenden Fachkräfte in Abhängigkeit von verschiedenen Faktoren unterschiedlich vom Feedback profitiert haben, wurden die Regressionsanalysen nur für die Feedbackgruppe berechnet. Für lineare Regressionen gelten nach Bühner und Ziegler (2017) folgende Voraussetzungen:

- linearer Zusammenhang zwischen Prädiktoren und Kriterium
- Varianz des Fehlers ist in allen Bereichen gleich (Homoskedastizität)
- Normalverteilung der Residuen
- keine Multikollinearität.

Zur Überprüfung der Linearität wurde jeweils ein „Lack-of-Fit F-Test“ für den Zusammenhang zwischen dem jeweiligen Kriterium und den einzelnen Prädiktoren durchgeführt. Da



Abweichungen von der Linearität getestet werden, ist ein nicht signifikantes Ergebnis zur Annahme der Linearität erforderlich. Im Falle eines nicht linearen Zusammenhangs wurden die entsprechenden Prädiktoren aus den weiteren Analysen ausgeschlossen. In Anhang C7 ist ersichtlich zwischen welchen Variablen kein linearer Zusammenhang besteht. Um Multikollinearität auszuschließen, wurden die Kennwerte „VIF“ und „Toleranz“ geprüft. Ein „VIF“ über zehn und eine „Toleranz“ kleiner als 0.10 werden als kritisch betrachtet (Bühner & Ziegler, 2017). In keinem Regressionsmodell lag Multikollinearität der Prädiktoren vor. Das Vorliegen einer Normalverteilung der Residuen wurde mithilfe der Normalverteilungsdigramme eingeschätzt. Für die Beurteilung der Homoskedastizität wurde jeweils das Streudiagramm zwischen dem standardisierten vorausgesagten Kriterium und den standardisierten Residuen herangezogen. Bis auf einzelne Ausreißer in den Werten kann bei einem Großteil von Normalverteilung und Homoskedastizität ausgegangen werden. Heteroskedastizität beeinflusst nicht die Schätzung der Regressionsgewichte oder Determinationskoeffizienten, sondern die Signifikanzprüfung. In diesen Fällen wurde bei signifikanten Vorhersagen durch einen Prädiktor eine korrigierte Überschreitungswahrscheinlichkeit verwendet, die mithilfe des Bootstrap-Verfahrens ermittelt wurde.

In den Regressionsanalysen wurden jeweils im ersten Block die korrelierenden Störvariablen und im zweiten Hierarchieschritt zusätzlich die interessierenden Einflussvariablen als Prädiktoren betrachtet. Für jedes Kriterium wurde daher zunächst untersucht, zu welchen Störvariablen (z.B. Leistung zu T1 oder personenbezogene Variablen) ein signifikanter Zusammenhang besteht (siehe Anhang C8). Nur die Störvariablen, welche mit dem Differenzwert der jeweiligen Variable der sprachlichen Anregung signifikant zusammenhängen, wurden bei den Regressionsanalysen berücksichtigt. Für die kategorialen Variablen „Beruflicher Abschluss“, „Migrationshintergrund“ und „Funktion in der Einrichtung“ wurde eine Dummy-Kodierung vorgenommen. Dabei diente jeweils eine Kategorie als Bezugskategorie – für alle anderen Antwortkategorien wurden neue dichotome Dummy-Variablen gebildet (0 = Nein, 1 = Ja), welche dann bei den Regressionsanalysen berücksichtigt wurden. Inwiefern signifikante Effekte als praktisch bedeutsam einzustufen sind, kann mithilfe der Effektstärke  $R^2$  beurteilt werden, welche nach Bühner und Ziegler (2017) wie folgt interpretiert werden kann: kleiner Effekt:  $R^2 = 0.02$ , mittlerer Effekt:  $R^2 = 0.13$ , großer Effekt:  $R^2 = 0.26$ .

Da in erster Linie von Interesse ist, wieviel zusätzliche Varianz über die Störvariablen hinaus durch die jeweiligen potenziellen Einflussvariablen aufgeklärt werden kann, ist dabei vor allem die Änderung in  $R^2$  ( $\Delta R^2$ ) vom ersten zum zweiten Modell relevant.

## 7 Umsetzung und Überprüfung des Feedback-Konzepts: Ergebnisse

Das Ziel der vorliegenden Studie war die wissenschaftliche Überprüfung des Feedback-Konzepts hinsichtlich der Umsetzbarkeit und der Effekte auf das sprachliche Verhalten der Fachkräfte. Die Ergebnisse der Untersuchung werden in diesem Kapitel thematisiert.

### 7.1 Umsetzung der Feedbackgespräche

Bevor die Effekte des Feedbackgesprächs dargestellt werden, soll in den folgenden Abschnitten zunächst die Umsetzung der Feedbackgespräche beschrieben werden. Dabei wird sowohl auf technische als auch auf inhaltliche Aspekte eingegangen.

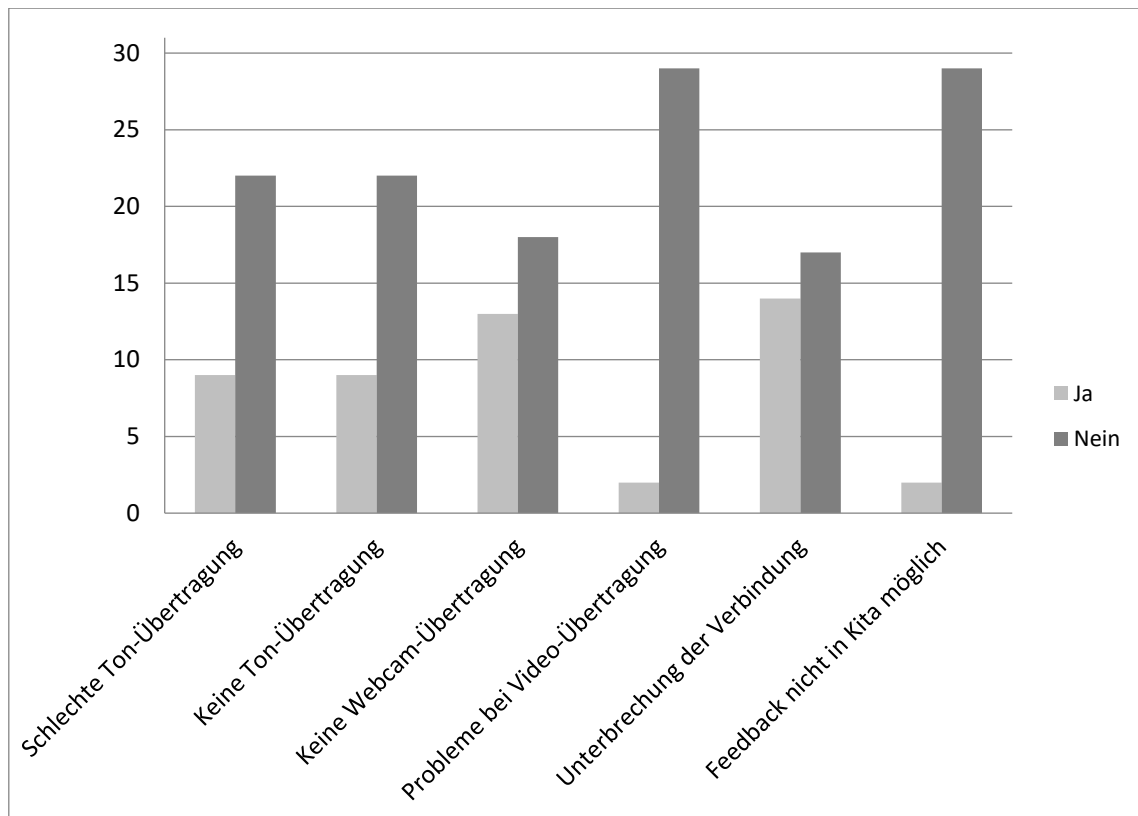
#### Technische Umsetzung

In Bezug auf die technische Umsetzung ist positiv anzumerken, dass alle Feedbackgespräche online mit der vitero-Software durchgeführt werden konnten. Wie Tabelle 18 zu entnehmen ist, sind bei fast zwei Drittel der 48 Gespräche allerdings technische Schwierigkeiten vorgekommen.

Tabelle 18: Häufigkeit von technischen Problemen während der Feedbackgespräche

		Häufigkeit	Prozent
Technische Probleme	Ja	31	64.6
	Nein	17	35.4

In Abbildung 10 ist dargestellt, wie häufig die verschiedenen Probleme bei den 31 Feedbackgesprächen aufgetaucht sind.



Abbildungung 10: Häufigkeit der verschiedenen technischen Probleme

Zu den häufigsten Problemen zählen Unterbrechungen der Sitzung ( $N = 14$ ) und eine nicht funktionierende Webcam-Übertragung ( $N = 13$ ). Die Feedbackgespräche waren zudem zum Teil durch eine schlechte Ton-Übertragung gekennzeichnet ( $N = 9$ ). Die Videosequenzen konnten in der Regel ohne Einschränkungen angeschaut werden ( $N = 2$ ). In zwei Fällen war es aufgrund der schlechten technischen Ausstattung oder fehlender Zugriffsrechte nicht möglich, die Software in der Kita zu installieren und das Feedback dort durchzuführen. Die Fachkräfte haben stattdessen von einem Computer zu Hause am Feedbackgespräch teilgenommen. Der Großteil der technischen Probleme ist auf eine zu schwache Internetverbindung in der Kita zurückzuführen.

### Ablauf und inhaltliche Umsetzung

Die Feedbackgespräche wurden nach dem in Kapitel 5.2 vorgestellten Konzept sowie der dazugehörigen Arbeitsvorlage durchgeführt. Der erste Schritt im Feedbackgespräch bestand immer darin, die Indikatoren für sprachförderliches Verhalten mit den Fachkräften zu besprechen. Im Anschluss daran wurden die vorher vom Feedbackgeber ausgewählten Indikatoren anhand der Videosequenzen reflektiert. Welche Indikatoren dabei wie häufig thematisiert wurden, ist in Tabelle 19 abgebildet.

Tabelle 19: In den Feedbackgesprächen thematisierte Indikatoren der Dimension „Language Modeling“

	Indikator 1		Indikator 2		Indikator 3	
	Häufigkeit	Prozent	Häufigkeit	Prozent	Häufigkeit	Prozent
Häufige Gespräche	4	8.3	15	31.3	27	56.3
Offene Fragen	3	6.3	31	64.6	14	29.2
Wiederholung u. Erweiterung	24	50	2	4.2	5	10.4
Handlungsbegl. Sprechen	1	2.1	0	0	0	0
Differenzierte Sprache	6	12.5	0	0	2	4.2
Kein Indikator von LM	10	20.8				

Als Erstes wurde immer ein Aspekt besprochen, der bereits gut umgesetzt wurde (siehe Kapitel 5.2). Bei der Hälfte der Feedbackgespräche war dies der Indikator „Wiederholung und Erweiterung“. In Circa einem Fünftel der Gespräche wurde als Beispiel für positive Umsetzung ein nicht sprachrelevanter Aspekt thematisiert, weil nicht genügend positive Verhaltensbeispiele aus der Dimension „Language Modeling“ vorhanden waren. Gemäß den Vorgaben im Feedback-Konzept wurden nur Indikatoren der Dimension „Language Modeling“ hinsichtlich ihres Optimierungspotenzials reflektiert. Dabei wurden vor allem die Gespräche mit den Kindern sowie das Stellen von offenen Fragen besprochen. Das handlungsbegleitende Sprechen ist, wie bereits thematisiert, bei der Bilderbuchbetrachtung nur von nebensächlicher Bedeutung, da es wenige Handlungen gibt, die sprachlich begleitet werden können. Dementsprechend war dieser Aspekt fast nie Thema bei der Reflexion der Videosequenzen.

Auf Basis der besprochenen Inhalte wurden mit der Fachkraft schriftlich Ziele festgehalten. Aus den dabei entstandenen Zielvereinbarungen wurden die formulierten Ziele der Fachkräfte systematisch zusammengefasst. Tabelle 36 in Anhang C1 gibt einen Überblick über den Inhalt dieser Ziele. Das (mit Abstand) am häufigsten formulierte Ziel war das Stellen von offenen Fragen ( $N = 38$ ). Ca. ein Drittel der Fachkräfte hat sich vorgenommen, mehr mit den Kindern in den Dialog zu gehen ( $N = 16$ ), auf die Äußerungen der Kinder einzugehen/daran anzuknüpfen ( $N = 15$ ) sowie weniger vorwegzunehmen und mehr die Kinder sprechen zu lassen ( $N = 14$ ). Jeweils 12 Fachkräfte haben als Ziel angegeben, insgesamt weniger vorzulesen und mehr die Kinder erzählen lassen, den Kindern dabei mehr Zeit zum Antworten/Sprechen zu lassen und allgemein die Kinder stärker einzubeziehen. Von neun Fachkräften wurde das Wiederholen und Erweitern der kindlichen Äußerungen als Ziel geäußert. Darüber hinaus gab es eine Reihe von Zielen, die von wenigen Fachkräften formuliert wurden (z.B. Wörter/Begriffe erklären bzw. gemeinsam besprechen, Bezug zum Alltag

der Kinder herstellen, langsamer sprechen). Im Durchschnitt hat ein Feedbackgespräch ca. eine Stunde gedauert.

## 7.2 Bewertung des Feedbackgesprächs durch die Fachkräfte

Die am Feedback teilnehmenden Fachkräfte bewerteten zu T2 das Feedbackgespräch auf den fünf Ebenen „Satisfaction“, „Utility“, „Knowledge“, „Application to practice“ und „Individual organizational results“, wobei die Items so formuliert waren, dass ein Wert von 0 eine negative Bewertung und ein Wert von 10 eine positive Beurteilung bedeutet. Die deskriptiven Ergebnisse des Q4TE sind in Tabelle 20 zusammengefasst.

Tabelle 20: Bewertung des Feedbackgesprächs durch die Fachkräfte: Deskriptive Statistik für die fünf Skalen des Q4TE

	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>N</i>
Satisfaction	9.33	1.58	0.50	10	40
Utility	8.59	2.34	0	10	40
Knowledge	7.04	2.77	0	10	40
Application to practice	7.81	2.12	0.50	10	39
Individual organizational results	6.94	2.97	0	10	40

Die deskriptive Auswertung zeigt, dass die Fachkräfte insgesamt sehr zufrieden mit dem Feedback waren. Auch den Nutzen des Feedbackgesprächs für ihre pädagogische Arbeit schätzten die befragten Fachkräfte in der Regel sehr hoch ein. Im Durchschnitt stimmten sie zu ca. 70 % zu, dass sie durch das Feedbackgespräch neues Wissen erlernt haben. Sie gaben zudem mit einer Zustimmung von fast 80 % an, die im Feedback erworbenen Kenntnisse in ihrer täglichen Arbeit anzuwenden. Eine etwas schlechtere, aber dennoch positive Bewertung erhielt das Feedback hinsichtlich der Effekte auf ihre allgemeine Arbeitsweise („Individual organizational results“).

Die Einschätzung mithilfe des Q4TE wurde durch freie Kommentare der Fachkräfte im Fragebogen ergänzt. Dazu liegen die Antworten von insgesamt 40 Fachkräften vor.

Auf die Frage, was sie aus dem Feedbackgespräch für ihre pädagogische Arbeit mitgenommen haben, antworteten die Fachkräfte mit Abstand am häufigsten, bei einer Bilderbuchbetrachtung mehr (offene) Fragen an die Kinder stellen zu wollen ( $N = 19$ ). Zum Teil wurde auch genannt, die Kinder allgemein mehr einzubeziehen und erzählen wollen zu lassen ( $N = 8$ ). Dabei scheint es den Fachkräften wichtig zu sein, den Kindern ausreichend Zeit zum Antworten bzw. Erzählen zu geben ( $N = 7$ ). Einige Fachkräfte haben Gesprächspotenzial gesehen, wenn sie noch mehr auf die Äußerungen der Kinder eingehen und deren

Beiträge aufgreifen ( $N = 5$ ). Vereinzelt wurde als Erkenntnis das Erklären von neuen Begriffen ( $N = 3$ ) und die Wiederholung kindlicher Äußerungen formuliert ( $N = 1$ ). Viele der Aspekte spiegeln demnach die Ziele wieder, welche am Ende des Feedbackgesprächs formuliert wurden. Neben neuen Anregungen hat ein Teil der Fachkräfte auch eine Art Bestätigung und Bestärkung aus dem Feedback gezogen ( $N = 6$ ).

Die meisten Fachkräfte haben angegeben, die bei der ersten Frage genannten Aspekte bereits umzusetzen. Dementsprechend war auch bei dieser Frage die häufigste Antwort, dass sie mehr (offene) Fragen stellen ( $N = 19$ ). Nach eigenen Angaben scheint es einigen zudem besser zu gelingen, die Kinder mehr erzählen zu lassen ( $N = 8$ ), den Kindern mehr Zeit zum Antworten zu geben ( $N = 5$ ) und noch häufiger auf die Beiträge der Kinder einzugehen ( $N = 4$ ).

Am meisten am Feedbackgespräch gefallen hat den teilnehmenden Fachkräften, dass sie mithilfe der Videosequenzen die Möglichkeit hatten, sich selbst und die (Reaktion der) Kinder einmal „von außen“ zu betrachten und das eigene Verhalten auf diese Weise reflektieren zu können ( $N = 21$ ). Auch die anderen Bausteine des Feedbacks in Form der zu Beginn besprochenen Indikatoren sowie der am Ende schriftlich festgehaltenen Zielsetzungen wurden von manchen Fachkräften ( $N = 8$ ) positiv wahrgenommen. Vielen Fachkräften hat zudem gefallen, dass im Feedbackgespräch sowohl Stärken als auch Entwicklungspotenzial thematisiert wurden. So wurden zum Beispiel teilweise Bestärkung und Lob als gelungene Aspekte des Feedbackgesprächs genannt ( $N = 8$ ). Gleichmaßen bedeutend war für einen Teil der Fachkräfte aber auch, dass während des Feedbacks gemeinsam erarbeitet wurde, wie man die entsprechenden Situationen noch sprachanregender gestalten könnte ( $N = 6$ ). In diesem Kontext wurde von manchen Fachkräften ( $N = 3$ ) die Art und Weise geschätzt, mit der das eigene Optimierungspotenzial aufgezeigt und erläutert wurde. Für einen großen Teil der Fachkräfte hat auch die angenehme Atmosphäre sowie die freundliche, offene und wertschätzende Kommunikation einen wesentlichen Beitrag zum Gelingen des Feedbackgesprächs beigetragen ( $N = 17$ ). Vereinzelt wurden die gute Vorbereitung ( $N = 1$ ), die technische Umsetzung ( $N = 3$ ) oder die Ausführlichkeit des Feedbacks ( $N = 3$ ) genannt.

In Bezug auf die Frage, was den Fachkräften nicht so gut gefallen hat, haben viele nichts angegeben bzw. geantwortet, dass ihnen alles gefallen hat. Ein kleiner Teil der Fachkräfte hat angegeben, dass die Technik besser hätte funktionieren können und die technische Ausstattung in der Kita nicht ausreichend war ( $N = 5$ ).

Dies spiegelt sich auch in den Antworten zur Bewertung der technischen Umsetzung wieder. Viele Fachkräfte haben das online-basierte Videofeedback als sehr gut und unkompliziert empfunden ( $N = 14$ ). Ebenso häufig wurde festgehalten, dass Online-Feedback eine gute Möglichkeit darstellt, wenn die Technik funktioniert. Die Analyse der Videosequenzen

wurde von einem Teil der Fachkräfte als sehr gut geeignet für die Reflexion des eigenen Verhaltens betrachtet ( $N = 6$ ). Als Vorteile der internet-basierten Umsetzung wurden Zeiterparnis, Effizienz sowie zeitliche Flexibilität genannt ( $N = 4$ ). Für manche Fachkräfte scheint die Vorbereitung (z.B. dass Soft- und Hardware zur Verfügung gestellt und installiert wurden) zum unkomplizierten Umgang beigetragen zu haben ( $N = 4$ ). Einzelne Fachkräfte hätten ein persönliches Feedback bevorzugt ( $N = 5$ ).

Zusammenfassend kann festgehalten werden, dass das Feedbackgespräch und dessen Nutzen von den teilnehmenden Fachkräften als sehr positiv bewertet wurde.

## 7.3 Effekte des Feedbackgesprächs auf das sprachliche Interaktionsverhalten

In diesem Kapitel wird beleuchtet, ob sich neben der positiven Bewertung der Fachkräfte auch tatsächliche Effekte des Feedbacks im Verhalten der Fachkräfte gezeigt haben. Wie in Kapitel 6.3.1 beschrieben, wurde das sprachförderliche Interaktionsverhalten der Fachkräfte mit der Dimension „Language Modeling“ aus dem Beobachtungsinstrument CLASS beurteilt. Als zusätzlicher Indikator für die sprachliche Anregung durch die Fachkräfte wurde das beschriebene Kodier-Schema zur Ermittlung der konkreten Anzahl an sprachlichen Verhaltensweisen während der Bilderbuchbetrachtung herangezogen. Die Angaben beziehen sich immer auf die komplette Dauer einer Bilderbuchbetrachtung, welche im Durchschnitt zu T1 eine Länge von ca. 19 Minuten hatte. Dabei lag der Mittelwert in der Feedbackgruppe ( $M = 17.96$ ,  $SD = 8.05$ ) etwas höher als in der Kontrollgruppe ( $M = 19.78$ ,  $SD = 7.34$ ), wobei dieser Unterschied nicht signifikant ist ( $t(94) = 1.16$ ,  $p = 0.25$ ). Die Länge der Bilderbuchbetrachtung bleibt in der Kontrollgruppe zu T2 ( $M = 19.14$ ,  $SD = 6.78$ ) in etwa gleich ( $t(47) = 0.62$ ,  $p = 0.54$ ). In der Feedbackgruppe hingegen dauert die Bilderbuchbetrachtung durchschnittlich zu T2 ( $M = 23.69$ ,  $SD = 10.02$ ) länger als zu T1 ( $t(47) = -5.50$ ,  $p < 0.01$ ).

### 7.3.1 Effekte auf das „Language Modeling“ der Fachkräfte

Bevor auf die Effekte des online-basierten Feedbackgesprächs eingegangen wird, sollen zunächst kurz deskriptive Ergebnisse der Auswertung der Videoaufnahmen mit der Dimension „Language Modeling“ der CLASS vorgestellt werden.

#### Einschätzung der sprachlichen Anregung mit der Dimension „Language Modeling“

Mit einem Mittelwert von 2.65 ( $SD = 1.12$ ) liegt die beobachtete sprachliche Anregung der pädagogischen Fachkräfte zu T1 ( $N = 96$ ) im unteren Bereich einer mittleren Qualität (3-5) auf der sieben-stufigen Skala und entspricht somit in etwa dem durchschnittlichen Niveau

in amerikanischen ( $M = 2.85$ ,  $SD = 0.73$ ) und deutschen Kindergärten ( $M = 2.73$ ,  $SD = 0.76$ ) (siehe Kapitel 6.3.1). Auffällig ist dabei, dass die in dieser Studie vorliegende Standardabweichung von 1.12 deutlich über der mittleren Standardabweichung der anderen Studien liegt.

Um einen Hinweis auf die Zuverlässigkeit der Bewertungen zu erhalten, wurde die prozentuale Übereinstimmung zwischen den Ratern am Beispiel von 36 Videoaufnahmen ermittelt. Bei 80.56 % der Ratings erzielten die Rater eine Übereinstimmung von  $\pm 1$  auf der siebenstufigen Skala. Dies liegt etwas unter dem im Manual angegebenen Wert von 84.8 %, entspricht aber der durch die Autoren der CLASS vorgegebenen Anforderung von 80 % Übereinstimmung (Pianta, La Paro et al., 2008).

### Veränderungen im „Language Modeling“ durch das Feedbackgespräch

In Hypothese 1 wird die Annahme formuliert, dass die Teilnahme am Feedbackgespräch einen positiven Einfluss auf das sprachliche Verhalten der pädagogischen Fachkräfte hat. Dies würde bedeuten, dass Fachkräfte in der Feedbackbedingung signifikant größere Verbesserungen im „Language Modeling“ zwischen T1 und T2 zeigen als Fachkräfte in der Kontrollbedingung. Der Effekt des Feedbackgesprächs auf das „Language Modeling“ der Fachkräfte wurde mit einer zweifaktoriellen Varianzanalyse mit Messwiederholung untersucht. Tabelle 21 zeigt die Mittelwerte und Standardabweichungen der beiden Gruppen zu T1 und T2.

Tabelle 21: Deskriptive Statistik der Variable "Language Modeling" unterschieden nach Bedingung und Messzeitpunkt

		Language Modeling		
		T1	T2	Gesamt
<b>Kontrollgruppe</b>	<i>M</i>	2.77	2.67	2.72
	<i>SD</i>	1.10	1.23	1.09
<b>Feedbackgruppe</b>	<i>M</i>	2.52	3.27	2.90
	<i>SD</i>	1.15	1.11	1.02
<b>Gesamt</b>	<i>M</i>	2.65	2.97	
	<i>SD</i>	1.12	1.20	

Anmerkung:  $N = 48$  pro Gruppe und Messzeitpunkt;  $M$  = Mittelwert,  $SD$  = Standardabweichung

Der erste Haupteffekt überprüft, ob sich die Mittelwerte zwischen den zwei Bedingungen der unabhängigen Variable (Feedback- oder Kontrollgruppe) voneinander unterscheiden. Die Kontrollgruppe schneidet im „Language Modeling“ insgesamt etwas schlechter ab als



die Feedbackgruppe. Dieser Unterschied ist jedoch nicht signifikant ( $F(1,94) = 0.68$ ,  $p = 0.41$ ,  $\eta_p^2 = 0.01$ ).

Der zweite Haupteffekt betrachtet die Mittelwertdifferenz zwischen T1 und T2 (unabhängig von der Gruppe). Die Teilnehmer der Studie haben sich insgesamt von T1 zu T2 signifikant verbessert ( $F(1,94) = 12.47$ ,  $p < 0.01$ ,  $\eta_p^2 = 0.12$ ).

Betrachtet man die Mittelwerte in Tabelle 21 fällt jedoch auf, dass sich die Kontrollgruppe von T1 zu T2 etwas verschlechtert hat. Im Gegensatz dazu sind die Werte im „Language Modeling“ in der Feedbackbedingung angestiegen. Der Interaktionseffekt testet, ob diese Veränderungen signifikant unterschiedlich sind und gibt so Aufschluss darüber, ob die Teilnahme am Feedbackgespräch einen positiven Einfluss auf das „Language Modeling“ der Fachkräfte hat. Die inferenzstatistischen Befunde zum Interaktionseffekt (Messwiederholung \* Bedingung) zeigen einen signifikanten Interaktionseffekt ( $F(1,94) = 21.81$ ,  $p < 0.01$ ). Es handelt sich dabei um einen großen Effekt ( $\eta_p^2 = 0.19$ ). In Abbildung 11 und Abbildung 12 ist die vorhandene Interaktion zwischen Bedingung und Messzeitpunkt veranschaulicht.

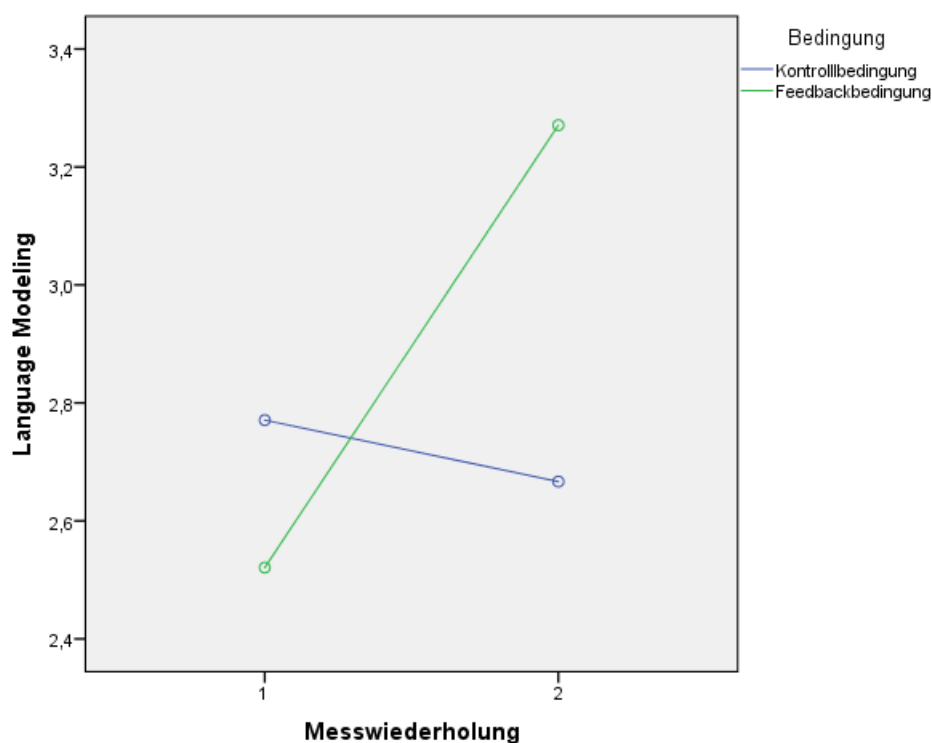


Abbildung 11: Profildigramm Haupteffekt Messwiederholung

In Abbildung 11 ist gut zu erkennen, dass die Rangordnung auf den beiden Linien vertauscht ist. Während in der Kontrollgruppe der Mittelwert zu T2 niedriger ist als zu T1, ist es

in der Feedbackgruppe genau umgekehrt. Insgesamt verbessern sich zwar im Durchschnitt die teilnehmenden Fachkräfte zu T2, dies scheint allerdings nicht für alle Fachkräfte gleichermaßen zu gelten, sondern damit zusammenzuhängen, an welcher Bedingung die jeweilige Fachkraft teilgenommen hat. Die Veränderung zwischen T1 und T2 (Haupteffekt Messwiederholung) lässt sich also nicht unabhängig von der Bedingung betrachten.

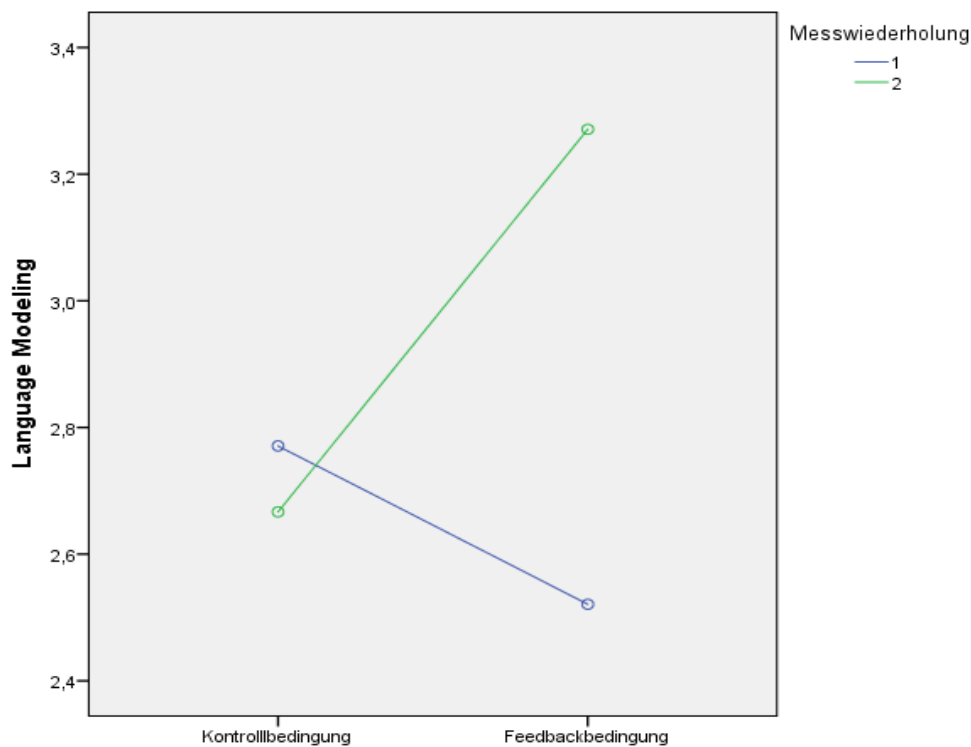


Abbildung 12: Profildigramm Haupteffekt Bedingung

Auch im Diagramm zum Haupteffekt der Bedingung (siehe Abbildung 12) überkreuzen sich beide Linien. Aus der vertauschten Rangordnung der Mittelwerte lässt sich ableiten, dass keine der beiden Bedingungen per se besser ist als die andere. Zu T1 weist die Kontrollgruppe etwas höhere Werte auf als die Feedbackgruppe, zu T2 hat die Feedbackgruppe hingegen deutlich höhere Werte als die Kontrollgruppe. Insgesamt sind die Unterschiede zwischen den beiden Bedingungen nicht von statistischer Bedeutsamkeit.

Es handelt sich demnach um eine disordinale Interaktion. Für keinen der beiden Haupteffekte lassen sich also Aussagen über die Rangfolge der Mittelwerte unabhängig von den Ausprägungen des anderen Faktors machen.

Die Ergebnisse der deskriptiven und varianzanalytischen Auswertung haben gezeigt, dass die pädagogischen Fachkräfte, die an einem online-basierten Feedbackgespräch teilgenommen haben, deutlich größere Verbesserungen in ihrem sprachanregenden Interakti-

onsverhalten zeigen als Fachkräfte, die kein Feedback erhalten haben. In einem zweiten Schritt wurde untersucht, wie groß die jeweiligen Unterschiede zwischen T1 und T2 für die beiden Gruppen sind. Dafür wurde sowohl für die Feedbackgruppe als auch für die Kontrollgruppe ein  $t$ -Test für abhängige Stichproben berechnet.

Die Kontrollgruppe hatte durchschnittlich zu T1 ( $M = 2.77$ ,  $SD = 1.10$ ) einen etwas höheren Wert im „Language Modeling“ als zu T2 ( $M = 2.67$ ,  $SD = 1.23$ ). Im  $t$ -Test für abhängige Stichproben verfehlt diese Differenz das Signifikanzniveau ( $t(47) = 0.87$ ,  $p = 0.39$ ,  $g = 0.13$ ). In der Feedbackbedingung verbesserten sich die Fachkräfte im Durchschnitt von 2.52 ( $SD = 1.15$ ) zu T1 auf 3.27 ( $SD = 1.11$ ) zu T2. Diese Verbesserung im „Language Modeling“ ist statistisch signifikant ( $t(47) = -5.43$ ,  $p < 0.01$ ) und entspricht einem großen Effekt ( $g = 0.78$ ).

Zusammenfassend kann festgehalten werden, dass sich die Fachkräfte in der Kontrollbedingung hinsichtlich ihres Wertes im „Language Modeling“ während der Bilderbuchbetrachtung von T1 zu T2 nicht bedeutend verändert haben, wohingegen in der Feedbackbedingung eine signifikante Verbesserung der Werte mit einer hohen Effektstärke festgestellt werden konnte. Hypothese 1 kann somit hinsichtlich der globalen Einschätzung der sprachlichen Anregung gestützt werden.

### 7.3.2 Effekte auf die Anzahl sprachlicher Verhaltensweisen

Inwiefern sich die positiven Effekte im „Language Modeling“ auch in der konkreten Anzahl an sprachlichen Verhaltensweisen wieder finden, soll an dieser Stelle thematisiert werden. Zunächst wird erneut auf die deskriptiven Ergebnisse der gesamten Stichprobe eingegangen.

#### Bewertung der Anzahl sprachlicher Verhaltensweisen

Wie viele der sprachlichen Verhaltensweisen im Durchschnitt von den pädagogischen Fachkräften während einer Bilderbuchbetrachtung gezeigt wurden, ist in Tabelle 22 ersichtlich. Für die deskriptiven Berechnungen wurden die Daten von T1 herangezogen, da diese den Ist-Stand unabhängig vom Feedbackgespräch darstellen.

Tabelle 22: Mittlere Anzahl an sprachlichen Verhaltensweisen während einer Bilderbuchbetrachtung für die Gesamtstichprobe zu T1 (N = 96)

Sprachliches Verhalten	<i>M</i>	<i>SD</i>	<i>Mi</i> <i>n</i>	<i>Max</i>
Längere Gespräche	0.13	0.44	0	3
Geschlossene Fragen mit Antwortmöglichkeit	13.31	10.35	0	43
Geschlossene Fragen ohne Antwortmöglichkeit	2.61	3.12	0	17
Ja/Nein-Fragen mit Antwortmöglichkeit	21.45	16.38	0	87
Ja/Nein-Fragen ohne Antwortmöglichkeit	3.67	5.54	0	27
Vermeintliche Fragen	2.77	4.04	0	16
Offene Fragen mit Antwortmöglichkeit	13.67	12.25	0	51
Offene Fragen ohne Antwortmöglichkeit	2.81	3.66	0	20
Aufforderungen zum Sprechen mit Antwortmöglichkeit	5.49	6.54	0	33
Aufforderungen zum Sprechen ohne Antwortmöglichkeit	0.45	0.91	0	6
Selftalk	0.65	1.40	0	7
Paralleltalk	0.96	2.68	0	20
Wiederholung ohne Korrektiv	14.78	12.53	0	73
Korrektive Wiederholung	1.06	1.70	0	9
Korrektive Erweiterung	0.69	1.48	0	11
Erweiterung ohne Korrektiv	7.74	6.64	0	34
Anknüpfen	4.98	5.53	0	29

Betrachtet man die die Mittelwerte in Tabelle 22 fällt auf, dass längere Dialoge mit einzelnen Kindern im Durchschnitt während einer Bilderbuchbetrachtung fast gar nicht beobachtet wurden. Auch korrekatives Feedback in Form von Wiederholungen oder Erweiterungen sowie handlungsbegleitendes Sprechen kommen nur sehr selten vor. Im Vergleich der nicht korrektiven Formen erreichen die Fachkräfte bei der reinen Wiederholung der kindlichen Äußerung einen höheren Mittelwert als bei Erweiterungen dieser. Mit Abstand am häufigsten scheinen Ja/Nein-Fragen (mit Antwortmöglichkeit) eingesetzt zu werden. Von allen gestellten Fragen zu T1 (N = 5788) machen sie 35.57 % aus. Geschlossene Fragen (22.08 %) und offene Fragen mit Antwortmöglichkeit (22.67 %) stellen jeweils ca. ein weiteres Fünftel der gestellten Fragen dar. In Tabelle 37 (Anhang C2) sind die Anzahl der Fragen zu T1 sowie die prozentuale Aufteilung dargestellt. Insgesamt auffällig erscheint die große Streuung in den Daten, was sich häufig in einer sehr großen Spannweite und einer hohen Standardabweichung äußert. Da es sich bei den Variablen um ein neu entwickeltes Kodierschema handelt, gibt es keine direkten Vergleichswerte. Eine Einordnung in die bisherige Literatur zu Befunden ähnlicher Untersuchungen erfolgt in Kapitel 8.2.

Als wie reliabel die Einschätzung der einzelnen Kodes eingeschätzt werden kann, wurde für die vorliegende Studie mithilfe der Intraklassenkorrelation anhand einer Teilstichprobe überprüft (siehe Kapitel 6.6.1). Die ermittelten ICC-Koeffizienten für die einzelnen Variablen sind in Tabelle 23 dargestellt.

Tabelle 23: ICC-Koeffizienten für die einzelnen sprachlichen Verhaltensweisen

Sprachliches Verhalten	ICC-Koeffizient
Längere Gespräche	0.03
Geschlossene Fragen mit Antwortmöglichkeit	0.92
Geschlossene Fragen ohne Antwortmöglichkeit	0.70
Ja/Nein-Fragen mit Antwortmöglichkeit	0.98
Ja/Nein-Fragen ohne Antwortmöglichkeit	0.79
Vermeintliche Fragen	0.77
Offene Fragen mit Antwortmöglichkeit	0.85
Offene Fragen ohne Antwortmöglichkeit	0.34
Aufforderungen zum Sprechen mit Antwortmöglichkeit	0.92
Aufforderungen zum Sprechen ohne Antwortmöglichkeit	-0.04
Selftalk	0.41
Paralleltalk	0.59
Wiederholung ohne Korrektiv	0.95
Korrektive Wiederholung	0.81
Korrektive Erweiterung	0.88
Erweiterung ohne Korrektiv	0.76
Anknüpfen	0.68

In Anbetracht der Tatsache, dass eine Intraklassenkorrelation von 0.70 als gute Reliabilität angesehen wird, sind die in Tabelle 23 abgebildeten Ergebnisse wie folgt einzuordnen. Eine gute bis sehr gute Interraterreliabilität konnte für folgende Variablen festgestellt werden: „Geschlossene Fragen mit/ohne Antwortmöglichkeit“, „Ja/Nein-Fragen mit/ohne Antwortmöglichkeit“, „Vermeintliche Fragen“, „Offene Fragen mit Antwortmöglichkeit“, „Aufforderungen zum Sprechen mit Antwortmöglichkeit“, „Wiederholung mit/ohne Korrektiv“, „Erweiterung mit/ohne Korrektiv“, „Anknüpfen“. Hinsichtlich der Variable „Paralleltalk“ kann von einer mittleren Reliabilität ausgegangen werden. Bei einzelnen Kodes ergaben die Analysen Hinweise auf eine Reliabilität von 0 („Längere Gespräche“ und „Aufforderungen zum Sprechen ohne Antwortmöglichkeit“) oder eine sehr geringe Reliabilität („Offene Fragen ohne Antwortmöglichkeit“ und „Selftalk“).

Da in der vorliegenden Studie die Videoaufnahmen von T1 und T2 immer vom selben Rater bewertet wurden, können auch Items mit einer geringeren Interraterreliabilität in die weiteren Analysen einbezogen werden.

### Veränderungen in der Anzahl sprachlicher Verhaltensweisen durch das Feedbackgespräch

Entsprechend der ersten Hypothese wurde angenommen, dass sich die Verbesserung der sprachlichen Anregung auch in der konkreten Anzahl sprachlicher Verhaltensweisen widerspiegelt. Es wurde deshalb davon ausgegangen, dass Fachkräfte in der Feedbackbedingung eine signifikant größere Verbesserung in der Anzahl sprachförderlicher Verhaltensweisen aufweisen als Fachkräfte in der Kontrollbedingung. Unter Berücksichtigung der thematisierten Inhalte in den Feedbackgesprächen (siehe Kapitel 5 und Kapitel 7.1) sind für vermeintliche Fragen bzw. Fragen oder Erzählaufforderungen ohne Antwortmöglichkeit keine Veränderungen oder eher Verringerungen in der Anzahl zu erwarten. Da das handlungsbegleitende Sprechen im Feedback nur eine nebensächliche Rolle gespielt hat, ist auch hier nicht von positiven Effekten auszugehen. Das Feedback-Konzept ist so angelegt, dass jede Fachkraft bei ihrem individuellen Leistungsstand abgeholt wird (siehe Kapitel 5.2.1). Bei Fachkräften, die beispielsweise zu T1 fast ausschließlich vorgelesen haben, wurde daher prinzipiell der Einbezug der Kinder durch das Stellen von Fragen anvisiert. Dementsprechend kann eine Zunahme in allen Frageformen mit Antwortmöglichkeit erwartet werden. Diese Annahmen wurden mit zweifaktoriellen Varianzanalysen mit Messwiederholung für jede Variable des Kodier-Schemas überprüft. Die Ergebnisse werden im Folgenden vorgestellt. Um dies möglichst übersichtlich zu gestalten, werden die Ergebnisse schrittweise, unterteilt in inhaltlich zusammenpassende Aspekte, präsentiert.

#### *Effekte auf die Anzahl an Fragen*

Zunächst sollen die von den Fachkräften an die Kinder gestellten Fragen beleuchtet werden. In Tabelle 24 sind die deskriptiven Befunde zu den beiden Haupteffekten abgebildet.

Tabelle 24: Deskriptive Statistik zu den beiden Haupteffekten der gestellten Fragen

	Kontrollgruppe		Feedbackgruppe		T1		T2	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Geschlossene Fragen mit AM	13.74	9.30	16.43	10.71	13.31	10.35	16.85	13.52
Geschlossene Fragen ohne AM	3.14	4.08	2.89	3.35	2.61	3.12	3.41	5.82
Ja/Nein-Fragen mit AM	20.04	12.99	25.32	14.53	21.45	16.38	23.92	15.17
Ja/Nein-Fragen ohne AM	3.97	5.84	3.72	3.56	3.67	5.54	4.02	5.16
Vermeintliche Fragen	2.82	3.03	2.76	3.02	2.77	4.04	2.81	3.56
Offene Fragen mit AM	13.40	10.60	18.72	13.22	13.67	12.25	18.45	15.15
Offene Fragen ohne AM	2.74	2.61	3.99	3.89	2.81	3.66	3.92	4.18

Anmerkung: Alle Mittelwerte basieren auf  $N = 96$  Beobachtungen

In den deskriptiven Befunden zum Haupteffekt der Bedingung zeichnet sich ab, dass bei manchen Variablen die Feedbackgruppe höhere Werte hat als die Kontrollgruppe (z.B. „Geschlossene Fragen mit AM“), und bei manchen Fragetypen ist es umgekehrt. Nur bei den „Offenen Fragen mit AM“ gibt es einen signifikanten Unterschied zwischen den zwei Gruppen mit höheren Werten in der Feedbackbedingung ( $F(1,94) = 4.74$ ,  $p = 0.03$ ,  $\eta_p^2 = 0.05$ ). In Bezug auf die Variablen „Ja/Nein-Fragen mit AM“ und „Offene Fragen ohne AM“ verfehlen die Mittelwertdifferenzen zwischen den Gruppen nur knapp das Signifikanzniveau. In Bezug auf alle anderen Fragen konnte kein signifikanter Unterschied zwischen der Kontroll- und der Feedbackbedingung gefunden werden. Die entsprechenden Kennwerte zum Haupteffekt der Bedingung können in Tabelle 38 (Anhang C3) eingesehen werden. Im Vergleich der Mittelwerte zwischen den Messzeitpunkten sind alle Mittelwerte zu T2 höher als zu T1. Für folgende Variablen konnte ein signifikanter Haupteffekt der Messwiederholung nachgewiesen werden: „Geschlossene Fragen mit AM“ ( $F(1,94) = 7.34$ ,  $p = 0.01$ ,  $\eta_p^2 = 0.07$ ), „Offene Fragen mit AM“ ( $F(1,94) = 17.85$ ,  $p < 0.01$ ,  $\eta_p^2 = 0.16$ ) und „Offene Fragen ohne AM“ ( $F(1,94) = 7.69$ ,  $p = 0.01$ ,  $\eta_p^2 = 0.08$ ). Bei den restlichen Frageformen ist der Unterschied zwischen den Messzeitpunkten nicht von statistischer Bedeutsamkeit (siehe Tabelle 39, Anhang C4). Mit Blick auf den Interaktionseffekt zwischen Messzeitpunkt und Be-

dingung sind in Tabelle 25 die deskriptiven Kennwerte für die Anzahl an Fragen pro Bilderbuchbetrachtung zu T1 und T2 für beide Gruppen getrennt dargestellt.

Tabelle 25: Deskriptive Statistik der gestellten Fragen – unterschieden nach Bedingung und Messzeitpunkt

		T1		T2	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Geschlossene Fragen mit AM	Kontrollgruppe	13.71	10.13	13.77	11.22
	Feedbackgruppe	12.92	10.66	19.94	14.98
Geschlossene Fragen ohne AM	Kontrollgruppe	2.81	3.22	3.46	6.89
	Feedbackgruppe	2.42	3.05	3.35	4.57
Ja/Nein-Fragen mit AM	Kontrollgruppe	21.06	13.74	19.02	14.20
	Feedbackgruppe	21.83	18.79	28.81	14.65
Ja/Nein-Fragen ohne AM	Kontrollgruppe	3.92	6.28	4.02	6.21
	Feedbackgruppe	3.42	4.73	4.02	3.90
Vermeintliche Fragen	Kontrollgruppe	3.06	4.17	2.58	3.80
	Feedbackgruppe	2.48	3.93	3.04	3.33
Offene Fragen mit AM	Kontrollgruppe	14.19	11.75	12.60	11.30
	Feedbackgruppe	13.15	12.82	24.29	16.33
Offene Fragen ohne AM	Kontrollgruppe	2.85	3.83	2.63	2.35
	Feedbackgruppe	2.77	3.53	5.21	5.14

Anmerkung: Alle Mittelwerte basieren auf  $N = 48$  Beobachtungen

Mit Blick auf die Mittelwerte in Tabelle 25, stellt sich ein differenzierteres Bild dar. Bei einigen Frageformen hat die Kontrollgruppe zu T1 einen höheren Wert als die Feedbackgruppe. Dies ändert sich jedoch zum Teil zu T2, wo die Feedbackgruppe einen größeren Mittelwert aufweist (z.B. bei den geschlossenen Fragen mit Antwortmöglichkeit). Die inferenzstatistische Auswertung ergab für die folgenden Variablen einen signifikanten Interaktionseffekt: „Geschlossene Fragen mit AM“ ( $F(1,94) = 7.08, p = 0.01, \eta_p^2 = 0.07$ ), „Ja/Nein-Fragen mit AM“ ( $F(1,94) = 9.85, p < 0.01, \eta_p^2 = 0.10$ ), „Offene Fragen mit AM“ ( $F(1,94) = 31.63, p < 0.01, \eta_p^2 = 0.25$ ) sowie „Offene Fragen ohne AM“ ( $F(1,94) = 11.22, p < 0.01, \eta_p^2 = 0.11$ ). Die Veränderung zwischen T1 und T2 hinsichtlich dieser Fragetypen hängt demnach davon ab, in welcher Bedingung die einzelne Fachkraft war. Bei den geschlossenen Fragen ohne Antwortmöglichkeit ( $F(1,94) = 0.06, p = 0.80, \eta_p^2 = 0.00$ ), den Ja/Nein-Fragen ohne Antwortmöglichkeit ( $F(1,94) = 0.27, p = 0.60, \eta_p^2 = 0.00$ ) sowie den vermeintlichen Fragen ( $F(1,94) = 1.20, p = 0.28, \eta_p^2 = 0.01$ ) konnte kein statistisch bedeutsamer Interaktionseffekt festgestellt werden.



Die signifikanten Interaktionseffekte der genannten Variablen werden anhand der Profildiagramme und der Ergebnisse der  $t$ -Tests im Folgenden näher betrachtet.

Abbildung 13 zeigt, dass die Kontrollbedingung zu T1 im Durchschnitt etwas mehr geschlossene Fragen mit Antwortmöglichkeit gestellt hat als die Feedbackgruppe, zu T2 hingegen haben die Fachkräfte der Feedbackbedingung einen höheren Mittelwert. Unterschiede zwischen den Gruppen können daher nur in Zusammenhang mit dem Messzeitpunkt interpretiert werden. Im Gegensatz dazu gilt der Haupteffekt für den Messzeitpunkt auch unabhängig von der Bedingung – die Anzahl der geschlossenen Fragen mit Antwortmöglichkeit ist in beiden Gruppen zu T2 höher als zu T1 (siehe Abbildung 14). Die Interaktionsart wird in diesem Fall als hybrid bezeichnet. Obwohl sich beide Gruppen in Bezug auf die Variable „Geschlossene Fragen mit AM“ verbessert haben, so ist auch in Abbildung 14 ein Unterschied in der Stärke der Veränderung zwischen den Gruppen erkennbar. Bei der Kontrollgruppe ergab der  $t$ -Test für abhängige Stichproben, dass der Unterschied zwischen T1 und T2 nicht signifikant ist ( $t(47) = -0.04$ ,  $p = 0.97$ ,  $g = 0.01$ ). In der Feedbackbedingung haben die Fachkräfte hingegen zu T2 bedeutend mehr geschlossene Fragen mit Antwortmöglichkeit gestellt als zu T1 ( $t(47) = -3.30$ ,  $p < 0.01$ ). Diese Differenz entspricht einem mittleren Effekt ( $g = 0.48$ ).

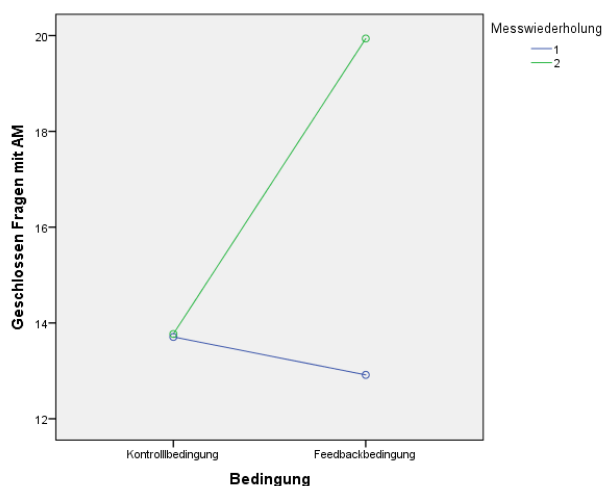


Abbildung 13: Profildiagramm Haupteffekt Bedingung – Geschlossene Fragen mit AM

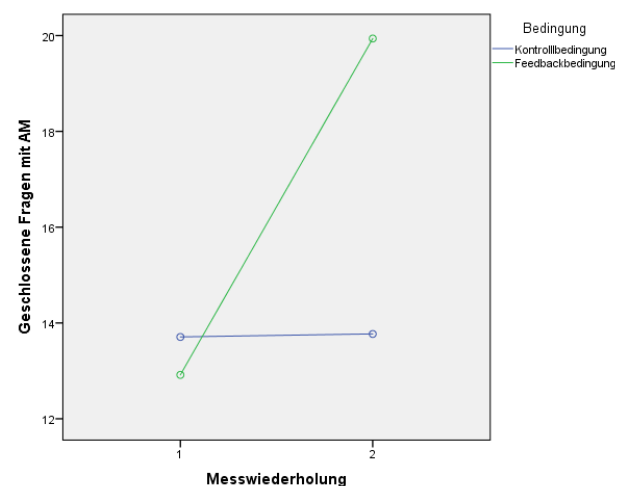


Abbildung 14: Profildiagramm Haupteffekt Messzeitpunkt – Geschlossene Fragen mit AM

In Abbildung 15 ist zu erkennen, dass die Feedbackgruppe zu beiden Messzeitpunkten höhere Werte in der Variable „Ja/Nein-Fragen mit AM“ hat. Der Haupteffekt der Bedingung ist daher unabhängig vom Messzeitpunkt. Wie bereits beschrieben, verfehlt dieser Haupteffekt allerdings knapp das Signifikanzniveau. Beim Haupteffekt des Messzeitpunkts kommt es zu einer Umkehrung der Rangfolge. Wie in Abbildung 16 dargestellt ist, sind die

Werte der Kontrollgruppe zu T1 höher als zu T2. Diese Verschlechterung in den „Ja/Nein-Fragen mit AM“ ist jedoch nicht signifikant ( $t(47) = 1.38, p = 0.18, g = 0.20$ ). Die Fachkräfte in der Feedbackbedingung haben zu T2 signifikant mehr „Ja/Nein-Fragen mit AM“ gestellt als zu T1 ( $t(47) = -2.83, p = 0.01, g = 0.41$ ). Der Haupteffekt des Messzeitpunkts kann daher nur in Abhängigkeit von der Bedingung bewertet werden, weswegen es sich insgesamt erneut um eine hybride Interaktion handelt. Der Interaktionseffekt kann im Sinne eines positiven Effekts des Feedbackgesprächs interpretiert werden.

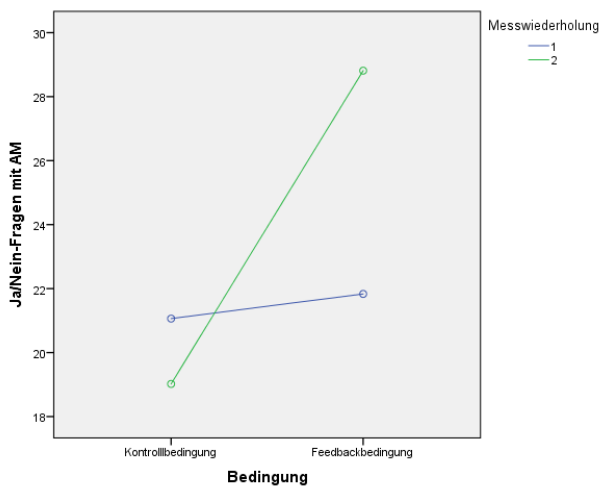


Abbildung 15: Profildiaqramm Haupteffekt Bedingung - Ja/Nein-Fragen mit AM

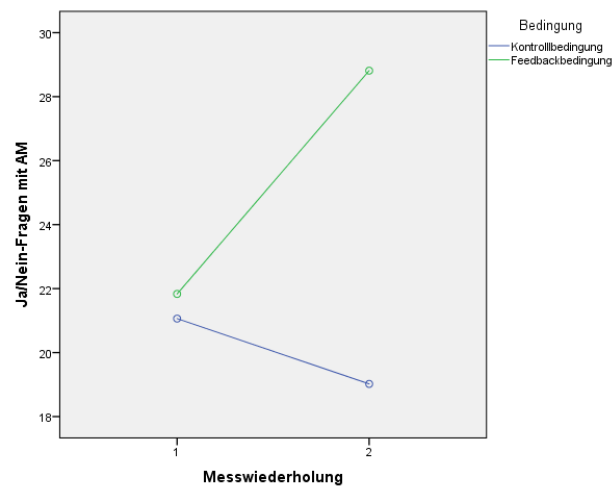


Abbildung 16: Profildiaqramm Haupteffekt Messzeitpunkt - Ja/Nein-Fragen mit AM

Bei der Variable „Offene Fragen mit AM“ gab es einen signifikanten Haupteffekt der Bedingung und des Messzeitpunktes. Sowohl in Abbildung 17 als auch in Abbildung 18 ist eine Umkehrung der Rangfolge der Mittelwerte festzustellen. Die Überlegenheit der Feedbackgruppe, welche im ersten Haupteffekt identifiziert wurde, gilt nur für den zweiten Messzeitpunkt. Zu T1 hatte die Kontrollgruppe etwas höhere Werte als die Feedbackgruppe. Der Haupteffekt des Messzeitpunktes hat auf eine allgemeine Verbesserung hinsichtlich der offenen Fragen mit Antwortmöglichkeit hingewiesen. Diese konnte allerdings nur für die Fachkräfte in der Feedbackbedingung bestätigt werden (siehe Abbildung 18). Die Differenz der Mittelwerte zwischen T1 und T2 entspricht in der Feedbackgruppe einem großen Effekt ( $t(47) = -6.04, p < 0.01, g = 0.87$ ). Die geringfügige Verschlechterung der Kontrollgruppe zu T2 hat keine statistische Bedeutsamkeit ( $t(47) = 1.21, p = 0.23, g = 0.17$ ). Gemäß dieser disordinalen Interaktion können die Effekte für die Bedingung und den Messzeitpunkt nur in gegenseitiger Abhängigkeit voneinander interpretiert werden.

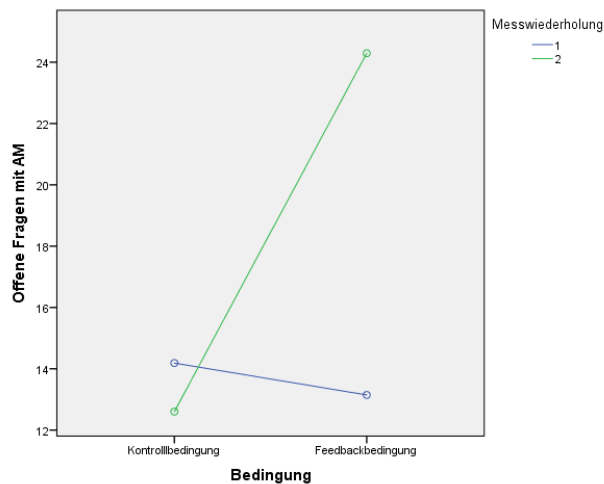


Abbildung 17: Profildiagramm Haupteffekt Bedingung – Offene Fragen mit AM

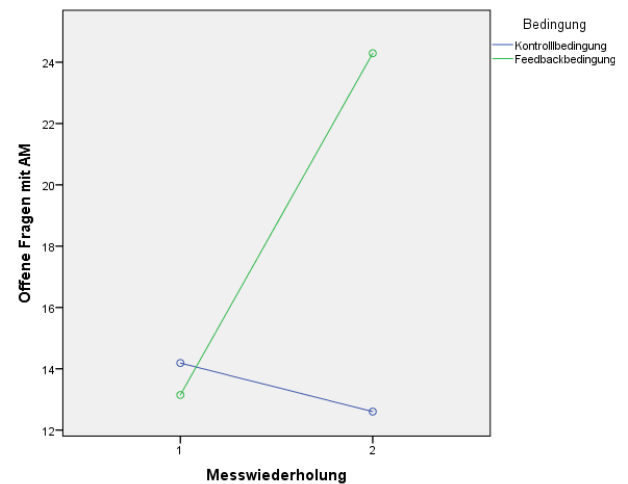


Abbildung 18: Profildiagramm Haupteffekt Messzeitpunkt – Offene Fragen mit AM

In Bezug auf die offenen Fragen ohne Antwortmöglichkeit zeigt sich ebenfalls eine disordiale Interaktion. Die Differenz zwischen den Gruppen, welche nur knapp das Signifikanzniveau verfehlt, ist je nach Messzeitpunkt unterschiedlich (siehe Abbildung 19). Der signifikante Haupteffekt der Messwiederholung hängt wiederum von der Bedingung ab (siehe Abbildung 20). Zu T1 konnte bei beiden Gruppen eine ähnliche Anzahl an offenen Fragen ohne Antwortmöglichkeit beobachtet werden. In der Kontrollbedingung hat diese zu T2 etwas abgenommen. Die Veränderung ist jedoch nicht signifikant ( $t(47) = 0.44$ ,  $p = 0.66$ ,  $g = 0.06$ ). Auf eine mittlere bis große Steigerung der Anzahl an offenen Fragen ohne Antwortmöglichkeit weist der  $t$ -Test für abhängige Stichproben für die Feedbackgruppe hin ( $t(47) = -4.05$ ,  $p < 0.01$ ,  $g = 0.59$ ).

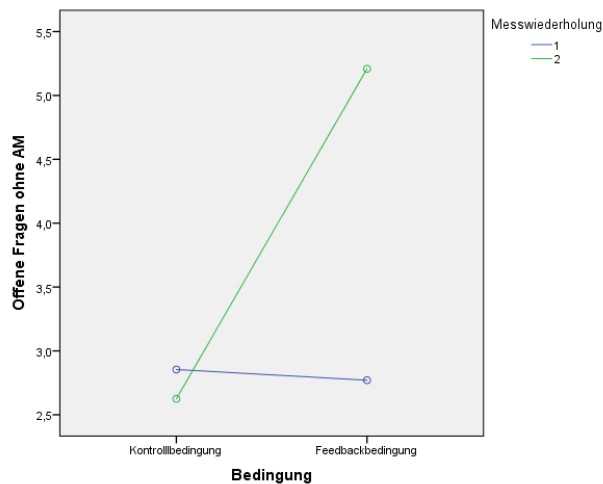


Abbildung 19: Profildigramm Haupteffekt Bedingung - Offene Fragen ohne AM

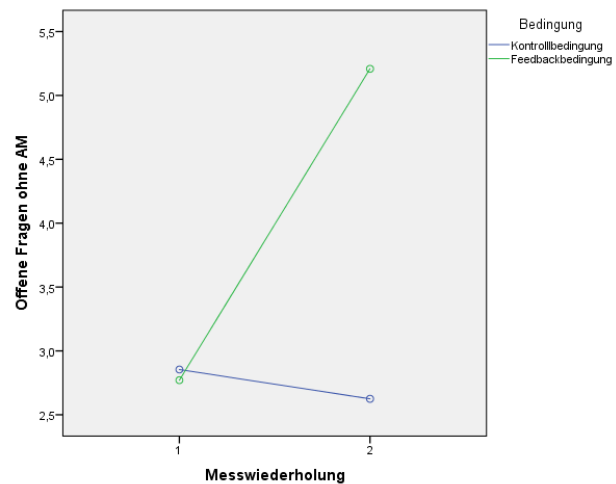


Abbildung 20: Profildigramm Haupteffekt Messzeitpunkt - Offene Fragen ohne AM

### Effekte auf die Anzahl an Erzählaufforderungen

Bei der Auswertung der Videoaufnahmen wurde nicht nur darauf geachtet, welche und wie viele Fragen die Fachkraft während der Bilderbuchbetrachtungen zu T1 und T2 gestellt hat, sondern auch inwiefern sie die Kinder ohne Fragen zum Sprechen aufgefordert hat. In Tabelle 26 sind die deskriptiven Befunde für die Variablen „Aufforderungen zum Sprechen mit Antwortmöglichkeit“ und „Aufforderungen zum Sprechen ohne Antwortmöglichkeit“ zusammengefasst.

Tabelle 26: Deskriptive Kennwerte für die Aufforderungen zum Sprechen mit und ohne Antwortmöglichkeit

Aufforderungen zum Sprechen							
		Mit Antwortmöglichkeit			Ohne Antwortmöglichkeit		
		T1	T2	Gesamt	T1	T2	Gesamt
Kontrollgruppe	M	6.23	6.38	6.30	0.58	0.31	0.45
	SD	7.45	7.21	6.26	1.05	0.72	0.78
Feedbackgruppe	M	4.75	10.29	7.52	0.31	0.54	0.43
	SD	5.46	15.12	9.01	0.72	1.15	0.78
Gesamt	M	5,49	8.33		0.45	0.43	
	SD	6.54	11.94		0.91	1.01	

Anmerkung: Die Mittelwerte pro Gruppe basieren auf  $N = 48$  Beobachtungen, die Gesamt-Mittelwerte beziehen sich auf ein  $N = 96$  Beobachtungen.

Wie in Tabelle 26 ersichtlich ist, hat die Feedbackgruppe insgesamt einen höheren Mittelwert hinsichtlich der Erzählaufforderungen mit Antwortmöglichkeit als die Kontrollgruppe.

Dieser Unterschied (Haupteffekt der Bedingung) ist nicht signifikant ( $F(1,94) = 0.59, p = 0.44, \eta_p^2 = 0.01$ ). Der Haupteffekt des Messzeitpunktes weist auf eine signifikante Verbesserung von T1 zu T2 hin ( $F(1,94) = 6.21, p = 0.01, \eta_p^2 = 0.06$ ). Vergleicht man die Mittelwerte der beiden Bedingungen, fällt auf, dass die Differenz zwischen T1 und T2 in der Feedbackgruppe deutlich höher ist als in der Kontrollgruppe. Dieser Unterschied entspricht einem signifikanten Interaktionseffekt mittlerer Stärke ( $F(1,94) = 5.59, p = 0.02, \eta_p^2 = 0.06$ ). In Abbildung 21 und Abbildung 22 wird deutlich, dass diese Interaktion hybrider Art ist. Der Unterschied zwischen den Gruppen ist nicht signifikant und lässt sich zudem nicht isoliert vom Messzeitpunkt interpretieren. Während die Kontrollgruppe zu T1 noch etwas besser ist als die Feedbackgruppe, kehrt sich dies zu T2 um (siehe Abbildung 21). Die Mittelwerte zu T2 sind, unabhängig von der Bedingung, im Durchschnitt höher als zu T1 (siehe Abbildung 22). Allerdings zeigt der  $t$ -Test für abhängige Stichproben, dass die Verbesserung von T1 zu T2 in der Kontrollgruppe nicht signifikant ist ( $t(47) = -0.13, p = 0.90, g = 0.02$ ). Die Fachkräfte in der Feedbackbedingung hingegen konnten sich signifikant verbessern ( $t(47) = -0.77, p = 0.01, g = 0.40$ ).

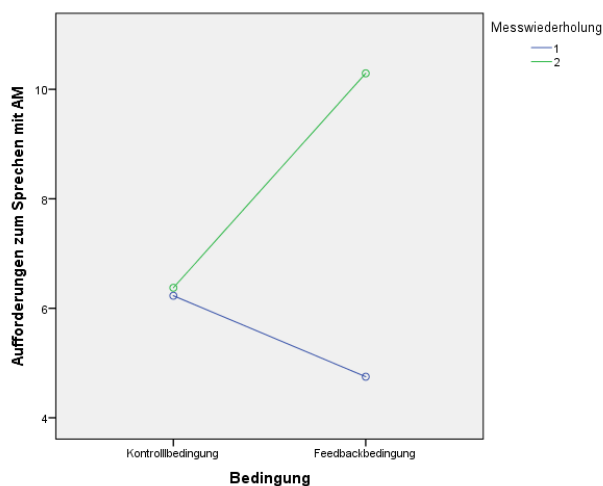


Abbildung 21: Profildigramm Haupteffekt Bedingung - Erzählaufforderungen mit AM

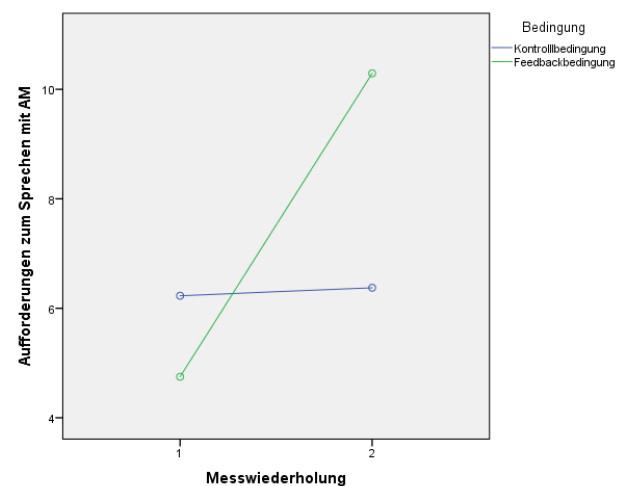


Abbildung 22: Profildigramm Haupteffekt Messzeitpunkt - Erzählaufforderungen mit AM

Mit einer Mittelwertdifferenz von 0.02 zwischen Kontroll- und Feedbackgruppe in der Variable „Aufforderungen zum Sprechen ohne AM“ (siehe Tabelle 26) kann kein signifikanter Haupteffekt der Bedingung festgestellt werden ( $F(1,94) = 0.02, p = 0.90, \eta_p^2 = 0.00$ ). Auch der Haupteffekt für den Messzeitpunkt verfehlt das Signifikanzniveau ( $F(1,94) = 0.03, p = 0.86, \eta_p^2 = 0.00$ ). Allerdings scheint es eine signifikante Interaktion zwischen Bedingung und Messzeitpunkt zu geben ( $F(1,94) = 4.86, p = 0.03, \eta_p^2 = 0.05$ ). Das Verhältnis der Anzahl der Aufforderungen zum Sprechen ohne Antwortmöglichkeit zu T1 kehrt sich zu T2 fast iden-

tisch um (siehe Tabelle 26). Diese disordinale Interaktion wird auch in Abbildung 23 und Abbildung 24 veranschaulicht.

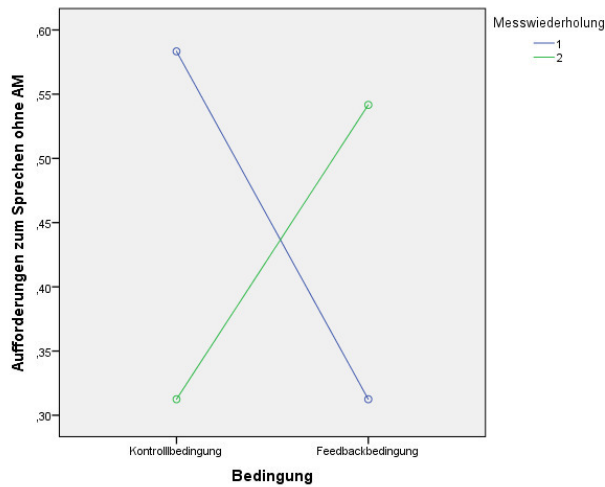


Abbildung 23: Profildiagramm Haupteffekt Bedingung - Erzählaufforderungen ohne AM

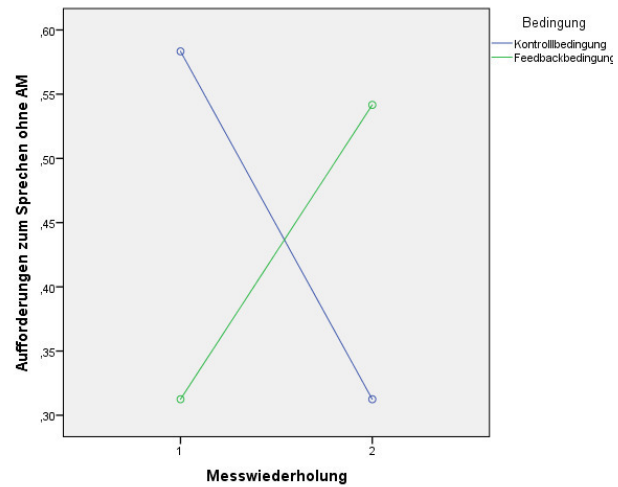


Abbildung 24: Profildiagramm Haupteffekt Messzeitpunkt - Erzählaufforderungen ohne AM

Der Interaktionseffekt weist darauf hin, dass sich die Veränderung zwischen T1 und T2 in beiden Gruppen unterschiedlich gestaltet. Während in der Kontrollgruppe die Werte abnehmen, steigen die der Feedbackgruppe an. Die Ergebnisse der  $t$ -Tests für abhängige Stichproben zeigen jedoch, dass die Differenz zwischen T1 und T2 sowohl in der Kontrollgruppe ( $t(47) = 1.70, p = 0.10, g = 0.25$ ) als auch in der Feedbackgruppe ( $t(47) = -1.42, p = 0.16, g = 0.21$ ) von geringer Stärke ist und das statistische Signifikanzniveau verfehlt.

#### *Effekte auf die Anzahl anknüpfender Äußerungen*

Wie viele Kommentare die Fachkraft macht, die sich inhaltlich auf die Äußerungen der Kinder beziehen, wurde mit der Variable „Anknüpfen“ erfasst. Die diesbezügliche deskriptive Statistik befindet sich in Tabelle 27.

.

Tabelle 27: Deskriptive Statistik für die Variable "Anknüpfen"

		Anknüpfen		
		T1	T2	Gesamt
Kontrollgruppe	<i>M</i>	5.08	5.08	5.08
	<i>SD</i>	5.58	5.53	5.00
Feedbackgruppe	<i>M</i>	4.88	7.50	6.19
	<i>SD</i>	5.54	8.49	6.49
Gesamt	<i>M</i>	4.98	6.29	
	<i>SD</i>	5.53	7.23	

Anmerkung: Die Mittelwerte pro Gruppe basieren auf  $N = 48$  Beobachtungen, die Gesamt-Mittelwerte beziehen sich auf  $N = 96$  Beobachtungen.

Betrachtet man zunächst nur die Gesamt-Mittelwerte, zeichnet sich ein höherer Mittelwert für die Feedbackgruppe und für den zweiten Messzeitpunkt ab. Die varianzanalytische Prüfung zeigt, dass der Unterschied zwischen den Gruppen (Haupteffekt der Bedingung) nicht signifikant ist ( $F(1,94) = 0.87$ ,  $p = 0.35$ ,  $\eta_p^2 = 0.01$ ). Der Haupteffekt des Messzeitpunkts hingegen erzielt statistische Bedeutsamkeit ( $F(1,94) = 5.46$ ,  $p = 0.02$ ,  $\eta_p^2 = 0.06$ ). Ein Blick in Tabelle 27 macht deutlich, dass die Erhöhung der Anzahl in der Variable „Anknüpfen“ von T1 zu T2 im Durchschnitt nur in der Feedbackgruppe beobachtet werden konnte. Der Mittelwert der Kontrollgruppe hat sich nicht verändert. Dies äußert sich in einem signifikanten Interaktionseffekt mittlerer Stärke ( $F(1,94) = 5.46$ ,  $p = 0.02$ ,  $\eta_p^2 = 0.06$ ). Zur Analyse der Interaktion zwischen Bedingung und Messzeitpunkt werden zusätzlich zu den deskriptiven Werten in Tabelle 27 wieder die Profildigramme der Haupteffekte herangezogen (siehe Abbildung 25 und Abbildung 26). Hinsichtlich des Haupteffekts der Bedingung ist eine Umkehrung der Rangfolge der Mittelwerte festzustellen. Während zu T1 die Kontrollgruppe einen etwas höheren Mittelwert als die Feedbackgruppe hat, ist das Gegenteil zu T2 der Fall. Da der Mittelwert der Kontrollgruppe zu T1 und T2 gleich hoch ist, kann man in Bezug auf den Haupteffekt des Messzeitpunkts nicht von einer umgekehrten Rangordnung sprechen. Der Unterschied zwischen T1 und T2 lässt sich ausschließlich für die Feedbackgruppe feststellen<sup>7</sup>. Es handelt sich dabei um eine signifikante Verbesserung ( $t(47) = -2.97$ ,  $p = 0.01$ ,  $g = 0.43$ ). Beide Haupteffekte lassen sich daher nur in Abhängigkeit der jeweils anderen Variable interpretieren (disordinale Interaktion). Für die Variable „Anknüpfen“ kann somit von einem mittelgroßen positiven Effekt des Feedbackgesprächs ausgegangen werden.

<sup>7</sup> Aufgrund des konstanten Mittelwerts in der Kontrollgruppe kann auf den t-Test für abhängige Stichproben verzichtet werden.

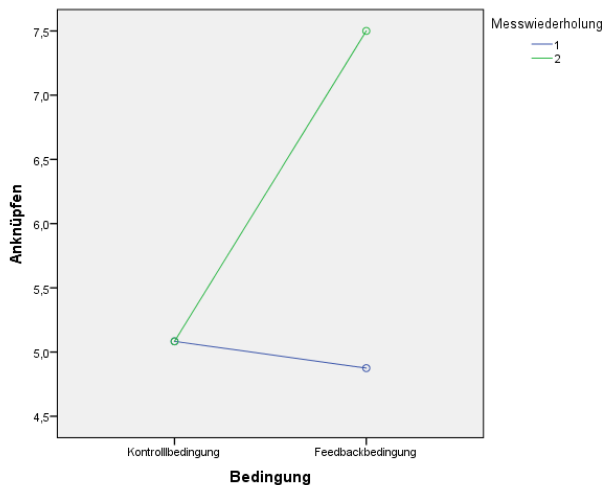


Abbildung 25: Profildigramm Haupteffekt Bedingung - Anknüpfen

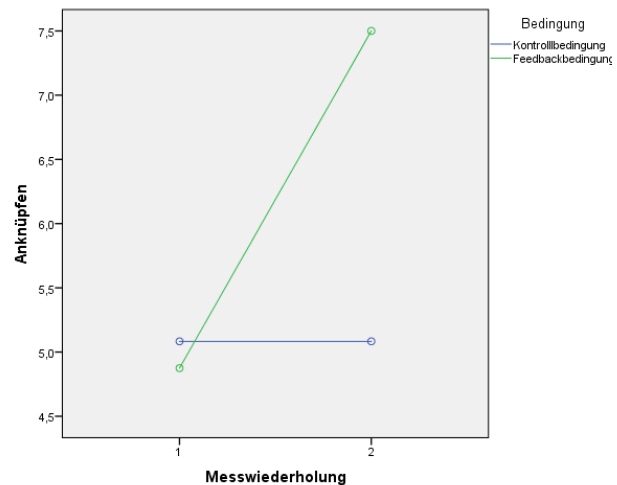


Abbildung 26: Profildigramm Haupteffekt Messzeitpunkt - Anknüpfen

### Effekte auf die Anzahl an längeren Dialogen

Mithilfe des Kodier-Schemas wurde auch erhoben, ob die Fachkraft durch Fragen, Erzählaufforderungen oder Kommentare längere Dialoge mit den Kindern aufbaut. Wie oft dies während einer Bilderbuchbetrachtung in den beiden Bedingungen zu den beiden Messzeitpunkten der Fall war, wird in Tabelle 28 dargestellt.

Tabelle 28: Deskriptive Kennwerte für die Anzahl der längeren Gespräche

		Längere Gespräche		
		T1	T2	Gesamt
Kontrollgruppe	<i>M</i>	0.08	0.23	0.16
	<i>SD</i>	0.35	0.52	0.29
Feedbackgruppe	<i>M</i>	0.17	0.46	0.31
	<i>SD</i>	0.52	0.80	0.46
Gesamt	<i>M</i>	0.13	0.34	
	<i>SD</i>	0.44	0.68	

Anmerkung: Die Mittelwerte pro Gruppe basieren auf  $N = 48$  Beobachtungen, die Gesamt-Mittelwerte beziehen sich auf  $N = 96$  Beobachtungen.

Insgesamt kamen längere Gespräche während der videografierten Bilderbuchbetrachtungen nur selten vor. Dabei wurden in der Feedbackgruppe insgesamt etwa doppelt so viele Gespräche verzeichnet wie in der Kontrollgruppe. Dieser Unterschied entspricht einem signifikanten Haupteffekt der Bedingung ( $F(1,94) = 3.97$ ,  $p = 0.05$ ,  $\eta_p^2 = 0.04$ ). Die Anzahl der längeren Gespräche hat sich insgesamt zwischen T1 und T2 von 0.13 auf 0.34 signifi-



kant erhöht ( $F(1,94) = 6.55, p = 0.01, \eta_p^2 = 0.07$ ). Es konnte kein signifikanter Interaktionseffekt festgestellt werden ( $F(1,94) = 0.73, p = 0.40, \eta_p^2 = 0.01$ ). Beide Haupteffekte können daher unabhängig interpretiert werden. Die Feedbackgruppe ist zu beiden Messzeitpunkten besser als die Kontrollgruppe und in beiden Gruppen kommt es zu einer Erhöhung der Anzahl an längeren Gesprächen. Der nicht signifikante Interaktionseffekt weist darauf hin, dass es in der Veränderung zwischen T1 und T2 keinen bedeutenden Unterschied zwischen Kontroll- und Feedbackbedingung gibt.

#### *Effekte auf das handlungsbegleitende Sprechen*

In Bezug auf das handlungsbegleitende Sprechen wurde zwischen „Selftalk“ und „Paralleltalk“ unterschieden. In Tabelle 29 sind die Mittelwerte und Standardabweichungen für diese beiden Variablen zusammengefasst.

Tabelle 29: Deskriptive Kennwerte für das handlungsbegleitende Sprechen

		Selftalk			Paralleltalk		
		T1	T2	Gesamt	T1	T2	Gesamt
<b>Kontrollgruppe</b>	<i>M</i>	0.83	0.79	0.81	1.02	0.92	0.97
	<i>SD</i>	1.69	1.20	1.15	2.36	2.03	1.89
<b>Feedbackgruppe</b>	<i>M</i>	0.46	0.79	0.63	0.90	1.10	1.00
	<i>SD</i>	1.01	1.41	1.02	2.98	2.79	2.78
<b>Gesamt</b>	<i>M</i>	0.65	0.79		0.96	1.01	
	<i>SD</i>	1.40	1.31		2.68	2.43	

Anmerkung: Die Mittelwerte pro Gruppe basieren auf  $N = 48$  Beobachtungen, die Gesamt-Mittelwerte beziehen sich auf  $N = 96$  Beobachtungen.

Wie die deskriptive Auswertung bereits andeutet, gibt es keinen signifikanten Unterschied zwischen den beiden Bedingungen in den Variablen „Selftalk“ ( $F(1,94) = 0.72, p = 0.40, \eta_p^2 = 0.01$ ) und „Paralleltalk“ ( $F(1,94) = 0.00, p = 0.95, \eta_p^2 = 0.00$ ). Zwischen T1 und T2 hat zwar in beiden Variablen eine Verbesserung stattgefunden, diese verfehlte jedoch sowohl bei „Selftalk“ ( $F(1,94) = 0.78, p = 0.38, \eta_p^2 = 0.01$ ) als auch bei „Paralleltalk“ ( $F(1,94) = 0.70, p = 0.79, \eta_p^2 = 0.00$ ) das Signifikanzniveau. Eine Interaktion zwischen Messzeitpunkt und Bedingung konnte weder in der Variable „Selftalk“ ( $F(1,94) = 1.29, p = 0.26, \eta_p^2 = 0.01$ ) noch im „Paralleltalk“ ( $F(1,94) = 0.63, p = 0.43, \eta_p^2 = 0.01$ ) festgestellt werden.

#### *Effekte auf das Wiederholen und Erweitern der kindlichen Äußerungen*

In einigen Feedbackgesprächen wurde das Wiederholen und Erweitern kindlicher Äußerungen thematisiert. Welche Effekte hinsichtlich der dazu erfassten Variablen nachgewiesen

werden konnten, wird im Folgenden in den Fokus genommen. Die deskriptive Statistik für die jeweiligen Haupteffekte ist in Tabelle 30 angegeben.

Tabelle 30: Deskriptive Statistik für die Haupteffekte in Bezug auf die Wiederholung und Erweiterung kindlicher Äußerungen

	Kontrollgruppe		Feedbackgruppe		T1		T2	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Wiederholung mit Korrektiv	0.94	1.16	1.76	2.08	1.06	1.70	1.64	2.76
Wiederholung ohne Korrektiv	13.60	9.40	18.93	13.67	14.78	12.53	17.75	14.16
Erweiterung mit Korrektiv	0.71	1.17	1.49	2.78	0.69	1.48	1.51	3.32
Erweiterung ohne Korrektiv	7.91	5.47	10.56	7.39	7.74	6.64	10.73	8.68

Anmerkung: Alle Mittelwerte basieren auf  $N = 96$  Beobachtungen

An den Werten kann abgelesen werden, dass die Feedbackgruppe in allen vier Variablen einen höheren Mittelwert hat als die Kontrollgruppe. Dieser Unterschied war für die folgenden Variablen signifikant: „Wiederholung mit Korrektiv“ ( $F(1,94) = 5.74, p = 0.02, \eta_p^2 = 0.06$ ), „Wiederholung ohne Korrektiv“ ( $F(1,94) = 4.94, p = 0.03, \eta_p^2 = 0.05$ ) sowie „Erweiterung ohne Korrektiv“ ( $F(1,94) = 4.00, p = 0.05, \eta_p^2 = 0.04$ ). Hinsichtlich der Erweiterungen mit korrigierender Funktion ergab die Varianzanalyse keinen signifikanten Haupteffekt der Bedingung ( $F(1,94) = 3.22, p = 0.08, \eta_p^2 = 0.03$ ). Die Mittelwerte in Tabelle 30 deuten bei allen Aspekten auf eine Verbesserung der Werte von T1 zu T2 hin. Sowohl bei den Wiederholungen ohne Korrektiv ( $F(1,94) = 7.15, p = 0.01, \eta_p^2 = 0.07$ ) als auch bei den Erweiterungen mit Korrektiv ( $F(1,94) = 8.72, p < 0.01, \eta_p^2 = 0.09$ ) und ohne Korrektiv ( $F(1,94) = 16.19, p = 0.00, \eta_p^2 = 0.15$ ) konnte die Differenz als signifikanter Haupteffekt des Messzeitpunkts bestätigt werden. In Bezug auf die Variable „Wiederholung mit Korrektiv“ entspricht die Veränderung einem kleinen bis mittleren Effekt, verfehlte jedoch knapp das Signifikanzniveau ( $F(1,94) = 3.64, p = 0.06, \eta_p^2 = 0.04$ ). Zur Überprüfung der Effekte des Feedbackgesprächs wurde darüber hinaus die Interaktion zwischen Messzeitpunkt und Bedingung analysiert. Tabelle 31 gibt einen Überblick über die deskriptiven Kennwerte für die beiden Bedingungen zu Messzeitpunkt 1 und 2.

Tabelle 31: Deskriptive Kennwerte für die Wiederholung und Erweiterung kindlicher Äußerungen – unterschieden nach Bedingung und Messzeitpunkt

		T1		T2	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Wiederholung mit Korrektiv	Kontrollgruppe	1.02	1.78	0.85	1.20
	Feedbackgruppe	1.10	1.64	2.42	3.56
Wiederholung ohne Korrektiv	Kontrollgruppe	14.58	10.50	12.63	9.80
	Feedbackgruppe	14.98	14.38	22.88	15.99
Erweiterung mit Korrektiv	Kontrollgruppe	0.58	0.82	0.83	1.95
	Feedbackgruppe	0.79	1.92	2.19	4.18
Erweiterung ohne Korrektiv	Kontrollgruppe	8.10	6.55	7.71	6.09
	Feedbackgruppe	7.38	6.77	13.75	9.82

Anmerkung: Alle Mittelwerte basieren auf  $N = 48$  Beobachtungen

Die Mittelwerte zeigen, dass die Überlegenheit der Feedbackgruppe nicht immer zu allen Messzeitpunkten vorliegt und die Veränderung zwischen T1 und T2 sich zum Teil sehr unterschiedlich zwischen den Bedingungen gestaltet. Dementsprechend zeigte sich in den Analysen bei allen vier Variablen ein signifikanter Interaktionseffekt. Die dazugehörigen Kennwerte können anhand der Tabelle 32 nachempfunden werden. Hinsichtlich der Wiederholungen und Erweiterungen mit Korrektiv kann von einem Interaktionseffekt kleiner bis mittlerer Stärke ausgegangen werden. Bei denen ohne korrektive Funktion handelt es sich sogar um große Effekte. Aufgrund der Verteilung der Mittelwerte in Tabelle 31 können die Interaktionseffekte bei allen vier Variablen zugunsten der Feedbackbedingung interpretiert werden. Eine positive Wirkung des Feedbackgesprächs kann somit für alle Formen der Wiederholung und Erweiterung angenommen werden.

Tabelle 32: Inferenzstatistische Kennwerte für den Interaktionseffekt – Wiederholung und Erweiterung

	<i>F</i>	<i>p</i>	$\eta_p^2$
Wiederholung mit Korrektiv	6.07	0.02*	0.06
Wiederholung ohne Korrektiv	19.70	<0.01**	0.17
Erweiterung mit Korrektiv	4.23	0.04*	0.04
Erweiterung ohne Korrektiv	20.77	0.00**	0.18

Anmerkung:  $df(\text{Interaktion}) = 1$ ,  $df(\text{Fehler}) = 94$

Die einzelnen Interaktionseffekte werden im Folgenden auf Basis der deskriptiven Kennwerte, der Profildigramme und der Ergebnisse der *t*-Tests beschrieben.

In Tabelle 31 ist erkennbar, dass bei der Variable „Wiederholung mit Korrektiv“ die Mittelwerte in der Kontrollgruppe immer kleiner sind als in der Feedbackgruppe. Die Rangord-

nung des Haupteffekts der Bedingung ist immer gleich – unabhängig vom Messzeitpunkt (siehe Abbildung 27). Tabelle 31 und Abbildung 28 zeigen, dass die Anzahl an Wiederholungen mit Korrektiv in der Kontrollgruppe von T1 zu T2 abnimmt, wohingegen sie in der Feedbackgruppe steigt. Bei der negativen Veränderung des Mittelwerts in der Kontrollbedingung handelt es sich nicht um einen bedeutenden Effekt ( $t(47) = 0.59$ ,  $p = 0.56$ ,  $g = 0.09$ ). Die Fachkräfte in der Feedbackbedingung konnten sich hingegen signifikant verbessern ( $t(47) = -2.48$ ,  $p = 0.02$ ,  $g = 0.36$ ). Es handelt sich demnach um eine hybride Interaktion.

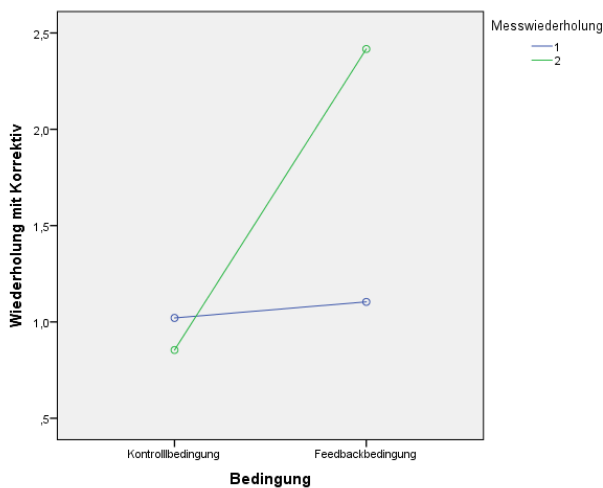


Abbildung 27: Profildiagramm Haupteffekt Bedingung – Wiederholung mit Korrektiv

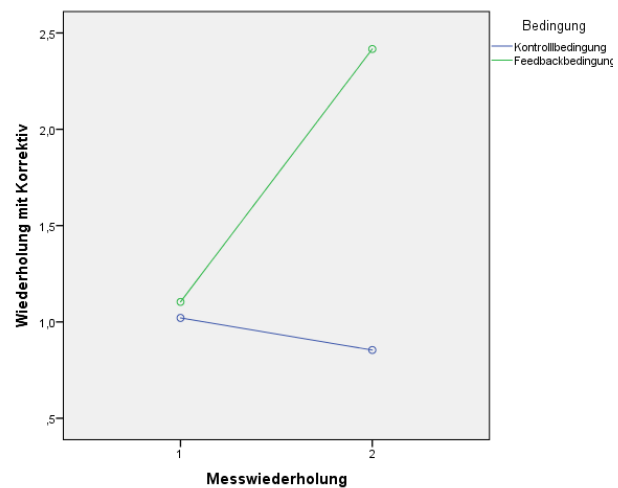


Abbildung 28: Profildiagramm Haupteffekt Messzeitpunkt – Wiederholung mit Korrektiv

Für die Variable „Wiederholung ohne Korrektiv“ zeigt sich ein ähnliches Bild. In der Kontrollgruppe konnten durchweg weniger Wiederholungen ohne Korrektiv während der Bilderbuchbetrachtung beobachtet werden. Die Unterschiede zwischen Kontroll- und Feedbackgruppe sind daher unabhängig vom Messzeitpunkt (siehe Abbildung 29). Der Haupteffekt des Messzeitpunkts kann nicht isoliert interpretiert werden. In Abbildung 30 ist eine Umkehrung der Rangfolge in den Mittelwerten zu erkennen. In der Kontrollbedingung ist der Mittelwert in der Anzahl an Wiederholungen ohne Korrektiv zu T1 höher als zu T2 – der Unterschied ist jedoch nur von geringer Stärke und verfehlt das Signifikanzniveau ( $t(47) = 1.76$ ,  $p = 0.08$ ,  $g = 0.26$ ). Bei den Fachkräften der Feedbackbedingung konnte eine signifikante Verbesserung festgestellt werden, die einem mittleren bis großen Effekt entspricht ( $t(47) = -4.11$ ,  $p < 0.01$ ,  $g = 0.59$ ).

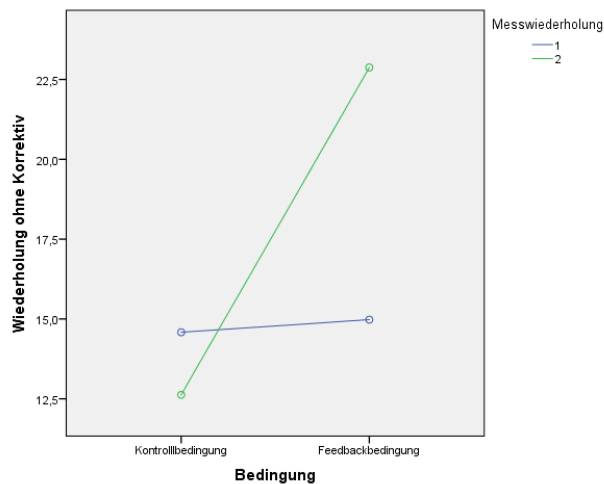


Abbildung 29: Profildigramm Haupteffekt Bedingung – Wiederholung ohne Korrektiv

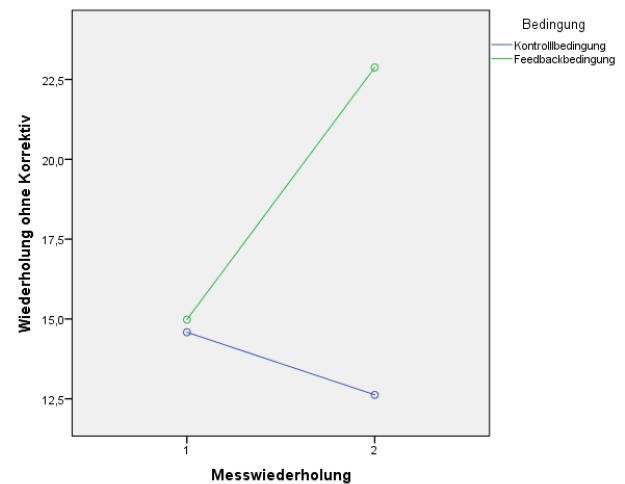


Abbildung 30: Profildigramm Haupteffekt Messzeitpunkt – Wiederholung ohne Korrektiv

Anhand der Mittelwerte in Tabelle 31 kann festgestellt werden, dass sich bezüglich der Erweiterungen mit Korrektiv beide Gruppen zwischen T1 und T2 verbessern, wobei die Feedbackgruppe zu beiden Messzeitpunkten höhere Werte hat als die Kontrollgruppe. Diese gleichbleibende Rangfolge der Mittelwerte wird in Abbildung 31 und Abbildung 32 veranschaulicht.

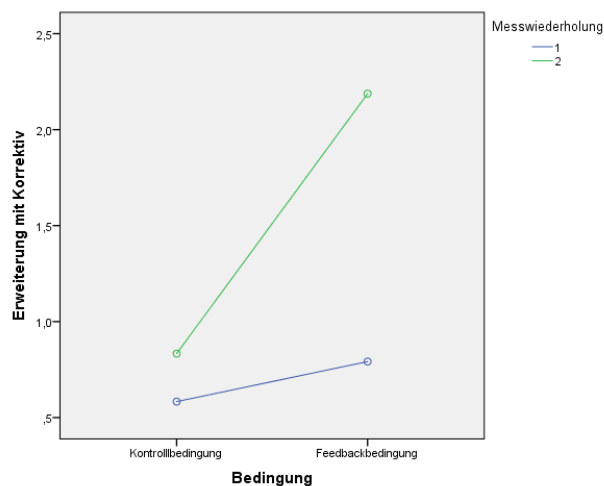


Abbildung 31: Profildigramm Haupteffekt Bedingung – Erweiterung mit Korrektiv

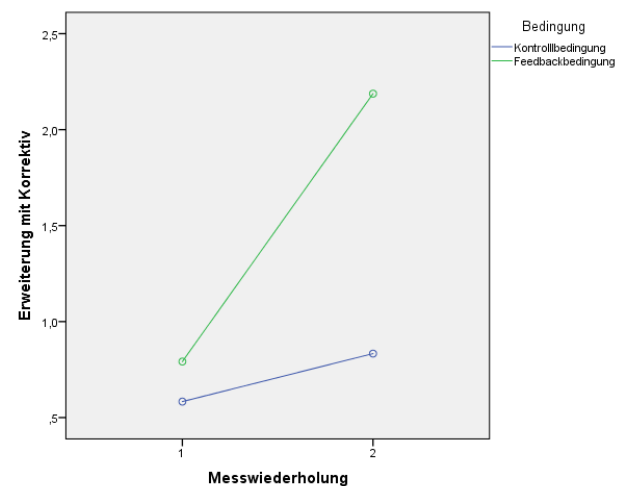


Abbildung 32: Profildigramm Haupteffekt Messzeitpunkt – Erweiterung mit Korrektiv

Sowohl der Haupteffekt der Bedingung als auch der Haupteffekt des Messzeitpunkts kann eindeutig interpretiert werden. Es handelt es sich also um eine ordinale Wechselwirkung zwischen Messzeitpunkt und Bedingung, bei der die Unterschiede zwischen den Messzeit-

punkten in einer Bedingung stärker ausfallen als in der anderen Bedingung. Die Ergebnisse der  $t$ -Tests für abhängige Stichproben geben Aufschluss darüber, wie groß die Mittelwertsdifferenz zwischen T1 und T2 in beiden Gruppen ist. In der Kontrollgruppe erzielt die Verbesserung von T1 zu T2 keine statistische Bedeutsamkeit ( $t(47) = -0.93$ ,  $p = 0.36$ ,  $g = 0.13$ ). Im Gegensatz dazu liegt bei den Fachkräften in der Feedbackgruppe eine signifikante Steigerung in der Anzahl an Erweiterungen ohne Korrektiv vor, welche als kleiner bis mittlerer Effekt angesehen werden kann ( $t(47) = -2.86$ ,  $p = 0.01$ ,  $g = 0.41$ ).

Während die Kontrollgruppe zu T1 einen höheren Mittelwert in der Variable „Erweiterung ohne Korrektiv“ hat als die Feedbackgruppe, kann zu T2 das Gegenteil beobachtet werden (siehe Tabelle 31). Es kann daher von einer umgekehrten Rangordnung der Mittelwerte in Bezug auf den Haupteffekt der Bedingung gesprochen werden (siehe Abbildung 33), weshalb dieser immer in Zusammenhang mit dem Messzeitpunkt bewertet werden sollte. Gleiches gilt für den Unterschied zwischen den Messzeitpunkten. Abbildung 34 zeigt, dass der Mittelwert in der Kontrollgruppe zu T2 abnimmt. Diese Veränderung ist allerdings nicht signifikant ( $t(47) = 0.43$ ,  $p = 0.67$ ,  $g = 0.06$ ). Die Ergebnisse des  $t$ -Tests für abhängige Stichproben belegen für die Feedbackgruppe eine signifikante Verbesserung in der Anzahl an Erweiterungen ohne Korrektiv in der Höhe eines großen Effekts ( $t(47) = -5.44$ ,  $p < 0.01$ ,  $g = 0.79$ ).

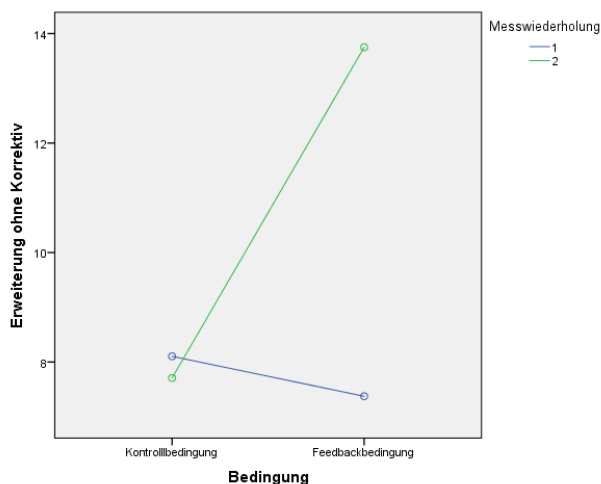


Abbildung 33: Profildiaagramm Haupteffekt Bedingung - Erweiterung ohne Korrektiv

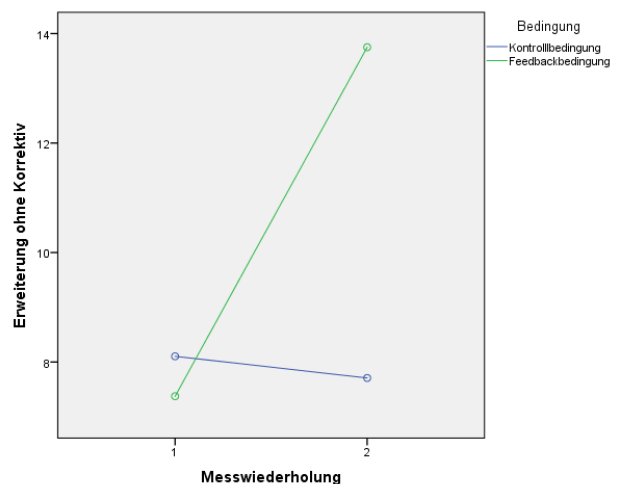


Abbildung 34: Profildiaagramm Haupteffekt Messzeitpunkt - Erweiterung ohne Korrektiv

Insgesamt betrachtet, konnten für die meisten der sprachlichen Verhaltensweisen signifikante Interaktionseffekte im Sinne einer positiven Wirkung des Feedbackgesprächs beobachtet werden. Tabelle 33 fasst noch einmal alle Interaktionseffekte sowie die dazugehö-

rigen Effektstärken zusammen. Variablen, bei denen eine signifikant positive Wirkung des Feedbacks festgestellt wurde, sind grau hinterlegt.

Tabelle 33: Signifikanz und Effektstärken der berechneten Interaktionseffekte für die Anzahl sprachlicher Verhaltensweisen

Sprachliches Verhalten	$p$	$\eta^2$
Längere Gespräche	0.40 (n.s.)	0.01
Geschlossene Fragen mit Antwortmöglichkeit	0.01**	0.07
Geschlossene Fragen ohne Antwortmöglichkeit	0.80 (n.s.)	0.00
Ja/Nein-Fragen mit Antwortmöglichkeit	<0.01**	0.10
Ja/Nein-Fragen ohne Antwortmöglichkeit	0.60 (n.s.)	0.00
Vermeintliche Fragen	0.28 (n.s.)	0.01
Offene Fragen mit Antwortmöglichkeit	<0.01**	0.25
Offene Fragen ohne Antwortmöglichkeit	<0.01**	0.11
Aufforderungen zum Sprechen mit Antwortmöglichkeit	0.02*	0.06
Aufforderungen zum Sprechen ohne Antwortmöglichkeit	0.03*	0.05
Selftalk	0.26 (n.s.)	0.01
Paralleltalk	0.43 (n.s.)	0.01
Wiederholung ohne Korrektiv	<0.01**	0.17
Korrektive Wiederholung	0.02*	0.06
Korrektive Erweiterung	0.04*	0.04
Erweiterung ohne Korrektiv	<0.01**	0.18
Anknüpfen	0.02*	0.06

Für alle Fragetypen und Erzählaufforderungen mit Antwortmöglichkeit liegen mittelgroße bis große Effekte des Feedbackgesprächs vor. Die größte Wirkung ist bei den offenen Fragen mit Antwortmöglichkeit zu verzeichnen. Die Anzahl an längeren Gesprächen ist zwar in der Feedbackgruppe etwas stärker angestiegen als in der Kontrollgruppe, dieser Unterschied ist jedoch nicht signifikant. Die Wirkung des Feedbackgesprächs auf die Variable „Anknüpfen“ kann als mittelstark betrachtet werden. Positive Effekte des Feedbackgesprächs zeigten sich zudem in allen Formen der Wiederholung und Erweiterung. Dabei sind die Effekte bei den Wiederholungen und Erweiterungen mit Korrektiv von geringer bis mittlerer Stärke. Die Veränderungen in den Wiederholungen und Erweiterungen ohne korrektive Funktion entsprechen großen Effekten. Wie bereits zu Beginn des Kapitels beschrieben, wurde für die vermeintlichen Fragen, die Fragen und Aufforderungen ohne Antwortmöglichkeit sowie das handlungsbegleitende Sprechen (Self- und Paralleltalk) kein Zuwachs angenommen. Dies konnte zum Teil bestätigt werden. Bei den geschlossenen Fragen und den Ja/Nein-Fragen ohne Antwortmöglichkeit wurde erwartungsgemäß kein

signifikanter Interaktionseffekt festgestellt. Die Anzahl an offenen Fragen und Erzählaufforderungen ohne Antwortmöglichkeit ist hingegen in der Feedbackbedingung stärker angestiegen als in der Kontrollbedingung. Hinsichtlich der Erzählaufforderungen ohne Antwortmöglichkeit ist allerdings anzumerken, dass die Ergebnisse des *t*-Tests gezeigt haben, dass weder in der Kontroll- noch in der Feedbackbedingung eine signifikante Veränderung stattgefunden hat. Es ist daher nicht von einer bedeutsamen Feedbackwirkung auszugehen<sup>8</sup>. In Bezug auf die Variablen „Selftalk“ und „Paralleltalk“ gab es, wie erwartet, keine Interaktion zwischen Messzeitpunkt und Bedingung und somit auch keine Effekte des Feedbackgesprächs.

Da bis auf wenige Ausnahmen eine positive Wirkung des Feedbackgesprächs ermittelt wurde, kann Hypothese 1 auch hinsichtlich der einzelnen sprachlichen Verhaltensweisen bestätigt werden.

### 7.3.3 Individuelle Veränderungen durch das Feedback

Mithilfe des „Reliable Change Index“ wurde analysiert, bei wie vielen Fachkräften durch das Feedback eine bedeutende Veränderung in der sprachlichen Anregung zwischen T1 und T2 stattgefunden hat. Dazu wurde anhand der in Kapitel 6.6.3 beschriebenen Formeln für jede Fachkraft aus der Feedbackgruppe ein „Reliable Change“-Wert ermittelt und geprüft, ob es sich mit einem Konfidenzintervall von 90 % bzw. 95 % um eine reliable Veränderung handelt. Tabelle 34 gibt einen Überblick darüber, bei wie vielen Fachkräften negative Veränderungen, keine Veränderungen bzw. positive Veränderungen festgestellt wurden. Mit negativen und positiven Veränderungen ist dabei keine Wertung verknüpft, sondern das Vorzeichen des Differenzwertes.

---

<sup>8</sup> Die Variable „Erzählaufforderungen ohne Antwortmöglichkeit“ ist daher trotz signifikanten Interaktionseffektes in Tabelle 33 nicht grau hinterlegt.



Tabelle 34: Häufigkeitsangaben zu Veränderungen nach dem "Reliable Change Index" von Jacobson und Truax (1991)

		5 %-Signifikanzniveau			10 %-Signifikanzniveau		
		-	/	+	-	/	+
Language Modeling	Häufigkeit	1	38	9	4	13	31
	Prozent	2.1	79.2	18.8	8.3	27.1	64.6
Längere Gespräche	Häufigkeit	1	42	5	1	42	5
	Prozent	2.1	87.5	10.4	2.1	87.5	10.4
Geschlossene Fragen mit AM	Häufigkeit	4	24	20	4	21	23
	Prozent	8.3	50	41.7	8.3	43.8	47.9
Geschlossene Fragen ohne AM	Häufigkeit	0	37	11	0	37	11
	Prozent	0	77.1	22.9	0	77.1	22.9
Ja/Nein-Fragen mit AM	Häufigkeit	8	13	27	8	12	28
	Prozent	16.7	27.1	56.3	16.7	25	58.3
Ja/Nein-Fragen ohne AM	Häufigkeit	0	43	5	0	33	15
	Prozent	0	89.6	10.4	0	68.8	31.3
Vermeintliche Fragen	Häufigkeit	3	41	4	4	35	9
	Prozent	6.3	85.4	8.3	8.3	72.9	18.8
Offene Fragen mit AM	Häufigkeit	1	30	17	2	22	24
	Prozent	2.1	62.5	35.4	4.2	45.8	50
Offene Fragen ohne AM	Häufigkeit	0	45	3	0	42	6
	Prozent	0	93.8	6.3	0	87.5	12.5
Aufforderungen zum Sprechen mit AM	Häufigkeit	5	33	10	5	30	13
	Prozent	10.4	68.8	20.8	10.4	62.5	27.1
Aufforderungen zum Sprechen ohne AM	Häufigkeit	0	44	4	0	44	4
	Prozent	0	91.7	8.3	0	91.7	8.3
Selftalk	Häufigkeit	0	45	3	0	45	3
	Prozent	0	93.8	6.3	0	93.8	6.3
Paralleltalk	Häufigkeit	0	46	2	0	46	2
	Prozent	0	95.8	4.2	0	95.8	4.2
Wiederholung ohne Korrektiv	Häufigkeit	3	21	24	4	19	25
	Prozent	6.3	43.8	50	8.3	39.6	52.1
Korrektive Wiederholung	Häufigkeit	2	37	9	4	29	15
	Prozent	4.2	77.1	18.8	8.3	60.4	31.3
Korrektive Erweiterung	Häufigkeit	1	36	11	1	36	11
	Prozent	2.1	75	22.9	2.1	75	22.9
Erweiterung ohne Korrektiv	Häufigkeit	1	33	14	1	29	18
	Prozent	2.1	68.8	29.2	2.1	60.4	37.5
Anknüpfen	Häufigkeit	1	38	9	1	38	9
	Prozent	2.1	79.2	18.8	2.1	79.2	18.8

Anmerkung: N = 48 (Feedbackgruppe); „-“: negative Veränderung, „/“: keine Veränderung, „+“: positive Veränderung

Eine Sichtung der Kennwerte zeigt, dass entsprechend der Ergebnisse der Varianzanalysen im handlungsbegleitenden Sprechen, den längeren Dialogen, den vermeintlichen Fragen sowie den Fragen ohne Antwortmöglichkeit verhältnismäßig weniger positive Veränderungen stattgefunden haben. Negative Veränderungen kommen bei den Fachkräften insgesamt nur selten vor. Verglichen mit den anderen Variablen ist die Anzahl an negativen Veränderungen in den Ja/Nein-Fragen mit Antwortmöglichkeit etwas höher.

In Bezug auf die globale Einschätzung des „Language Modeling“ konnten bei einem Konfidenzintervall von 95 % bei 18.8 % der Fachkräfte eine reliable positive Veränderung ermittelt werden. Mit einer Sicherheitswahrscheinlichkeit von 10 % zeigen sich hingegen bei fast zwei Drittel der Fachkräfte positive Veränderungen. Grund dafür ist, dass die Standardabweichung im „Language Modeling“ bei 1.12 liegt und entsprechend der Formel somit auf dem 5 %-Niveau alle Veränderungen kleiner als 1.17 nicht als reliabel angesehen werden. Dies betrifft daher alle Fachkräfte, die sich um einen Punkt auf der sieben-stufigen Skala verbessert haben.

Mit einer fünfprozentigen Fehlerwahrscheinlichkeit kann bei den Ja/Nein-Fragen mit Antwortmöglichkeit und den Wiederholungen ohne Korrektur von der größten Anzahl an positiven Veränderungen ausgegangen werden. Fast die Hälfte der Fachkräfte hat außerdem zu T2 bedeutend mehr geschlossene Fragen mit Antwortmöglichkeit gestellt.

Verglichen mit den großen Effektstärken in den Varianzanalysen, ist die Anzahl an positiven Veränderungen bei den offenen Fragen mit Antwortmöglichkeit auf dem 5 %-Niveau ( $N = 17$ ) relativ niedrig. Bei einem Konfidenzintervall von 90 % liegen deutlich mehr positive Veränderungen vor – hier weist die Hälfte der Fachkräfte Verbesserungen auf. Eine Sichtung der Daten zeigte, dass der *RC*-Wert vieler Fachkräfte bei den offenen Fragen mit Antwortmöglichkeit knapp unter 1.96 liegt und somit auf dem 5 %-Niveau die statistische Bedeutsamkeit verfehlt.

## 7.4 Einfluss von anderen Variablen auf die Effekte des Feedbacks

In Bezug auf die Rolle anderer Faktoren für die Wirksamkeit des Feedbacks wurden sowohl potenzielle Unterschiede zwischen Kontroll- und Feedbackbedingung als auch der Einfluss verschiedener personaler und situativer Merkmale innerhalb der Feedbackgruppe berücksichtigt.

### 7.4.1 Kontrolle von Störvariablen

Um sicherzustellen, dass die Effekte des Feedbackgesprächs nicht auf Unterschiede zwischen den beiden Gruppen in anderen Variablen zurückzuführen sind (Hypothese 2), wurde der Einfluss von potenziellen Störvariablen untersucht.

Dazu wurde zunächst mithilfe von Chi-Quadrat-Tests und *t*-Tests für unabhängige Stichproben geprüft, ob sich die Kontroll- und die Feedbackbedingung in relevanten Merkmalen signifikant unterscheiden (siehe Kapitel 6.6.3). Um den Fehler 2. Art zu vermeiden, wurde das Signifikanzniveau auf 20 % festgesetzt. Dabei wurden Merkmale der Fachkraft (Geschlecht, Alter, Migration und Mehrsprachigkeit, Qualifikation, Berufserfahrung, Position), situative Unterschiede zwischen T1 und T2 (Anzahl, Altersspanne und Sprachniveau der Kinder, Zeitabstand zwischen T1 und T2) sowie die Leistung und die Motivation zu T1 berücksichtigt.

Die Merkmale der Fachkraft wurden bereits in Kapitel 6.4 vergleichend für Kontroll- und Feedbackbedingung beschrieben. Tabelle 35 gibt einen Überblick über die situativen Rahmenbedingungen in den beiden Gruppen. Hinsichtlich der teilnehmenden Kinder sind die Rahmendbedingungen der Bilderbuchbetrachtungen zwischen Kontroll- und Feedbackgruppe vergleichbar. Im Durchschnitt haben jeweils zu T1 ca. sechs Kinder und zu T2 ca. sieben Kinder an der Bilderbuchbetrachtung teilgenommen. Dabei liegt die Altersspanne der Kinder zwischen 1,5 und 2 Jahren. Auch das Sprachniveau innerhalb der Kindergruppe ist sich sehr ähnlich. In beiden Gruppen wurde der Großteil der teilnehmenden Kinder von den Fachkräften als spracherfahren oder redegewandt eingeschätzt. Insgesamt haben die Erhebungen zu T1 und T2, wie geplant, ca. im Abstand von vier Wochen stattgefunden, wobei die Spannweite innerhalb der Feedbackgruppe größer ist.

Tabelle 35: Deskriptive Kennwerte zu den situativen Rahmenbedingungen bei der Bilderbuchbetrachtung

		Kontrollgruppe					Feedbackgruppe				
		<i>N</i>	<i>Min</i>	<i>Max</i>	<i>MW</i>	<i>SD</i>	<i>N</i>	<i>Min</i>	<i>Max</i>	<i>MW</i>	<i>SD</i>
Anzahl	T1	48	2	17	6.98	2.99	44	2	22	6.68	4.13
Kinder	T2	42	2	18	5.93	2.78	41	2	21	6.32	3.97
Altersspanne	T1	48	0	3,04	1.74	0.99	45	0	3,11	1.56	0.96
Kinder	T2	41	0	3,95	1.65	1.14	41	0	3,11	1.68	0.90
Anzahl Kinder, die...											
- (noch) nicht	T1	48	0	1	0.04	0.20	43	0	3	0.12	0.54
sprechen	T2	42	0	1	0.05	0.22	40	0	1	0.08	0.27
- nicht (Deutsch)	T1	48	0	2	0.21	0.46	43	0	1	0.16	0.37
sprechen mit DaZ	T2	42	0	3	0.17	0.58	40	0	1	0.10	0.30
- Sprachanfänger	T1	47	0	3	0.17	0.56	43	0	3	0.14	0.56
sind	T2	42	0	2	0.14	0.47	40	0	4	0.20	0.69
- Sprachanfänger	T1	47	0	4	0.34	0.82	43	0	4	0.37	0.79
sind mit DaZ	T2	42	0	3	0.26	0.63	40	0	4	0.30	0.79
- spracherfahren	T1	48	0	17	3.50	3.69	43	0	12	3.12	2.69
sind	T2	42	0	13	3.50	2.88	40	0	11	2.78	2.22
- redegewandt	T1	48	0	7	2.96	2.51	43	0	9	2.72	2.67
sind	T2	42	0	8	2.43	2.24	40	0	10	2.78	2.58
Abstand zwischen T1 und T2 in Wochen		48	3.57	4.86	4.10	0.29	48	2.71	7	4.23	0.98

Die Analysen ergaben signifikante Unterschiede hinsichtlich des Migrationshintergrunds der Fachkraft ( $\chi^2(1) = 5.04, p = 0.04, \phi = 0.23$ ), der allgemeinen Berufserfahrung als Fachkraft ( $t(90) = 1.87, p = 0.07, g = 0.39$ ) sowie der Differenz in der Anzahl der Kinder zwischen T1 und T2 ( $t(81) = 1.87, p = 0.07, g = 0.41$ ). In der Feedbackbedingung hatten demnach die Fachkräfte häufiger einen Migrationshintergrund sowie durchschnittlich weniger Berufserfahrung als Fachkraft. Ebenso war die Differenz in der Anzahl der Kinder zwischen T1 und T2 geringer, das heißt die Situationen waren in diesem Punkt vergleichbarer als in der Kontrollbedingung. In Bezug auf die übrigen Merkmale der Fachkräfte oder der Situationen konnten keine signifikanten Unterschiede gefunden werden. Die beiden Gruppen unterschieden sich zudem zu T1 in keiner der erfassten Variablen zur sprachlichen Anregung oder der Motivation. Alle Kennwerte können in Anhang C5 nachgeschlagen werden.

Der Einfluss des Migrationshintergrunds der Fachkraft wurde überprüft, indem die Variable als zusätzlicher Faktor in die Varianzanalyse mit Messwiederholung aufgenommen wurde. Dies erfolgte für alle Variablen der sprachlichen Anregung, bei denen ein signifikant positiver Effekt des Feedbacks festgestellt wurde (siehe Kapitel 7.3.1 und 7.3.2). Die diesbezüglichen

chen Kennwerte sind in Tabelle 42 in Anhang C6 dargestellt. Eine Sichtung der deskriptiven Kennwerte zeigt, dass Fachkräfte mit Migrationshintergrund tendenziell weniger sprachanregende Verhaltensweisen zeigen als Fachkräfte ohne Migrationshintergrund. Ein Einfluss auf die Effekte des Feedbacks konnte nicht festgestellt werden. Keiner der Interaktionseffekte zwischen Bedingung, Messwiederholung und Migrationshintergrund erreichte statistische Bedeutsamkeit. Sowohl bei Fachkräften mit Migrationshintergrund als auch bei denen ohne hat die Feedbackgruppe signifikant größere Verbesserungen in der sprachlichen Anregung als die Kontrollgruppe.

Zur Kontrolle des Einflusses der Berufserfahrung und der Anzahl der Kinder bei der Bilderbuchbetrachtung wurden bei allen abhängigen Variablen mit signifikantem Feedbackeffekt Kovarianzanalysen durchgeführt (siehe Tabelle 43 und Tabelle 44, Anhang C6). Die signifikanten Interaktionseffekte zwischen Bedingung und Messzeitpunkt bleiben unter Kontrolle der beiden Variablen bei allen Variablen der sprachlichen Anregung bestehen. Die Unterschiede in der Berufserfahrung und der Differenz in der Anzahl an Kindern bei der Bilderbuchbetrachtung zwischen Kontroll- und Feedbackgruppe haben demnach keine bedeutende Wirkung auf die Effekte des Feedbacks. Zusammenfassend kann daher festgehalten werden, dass die Feedbackeffekte auch unter Berücksichtigung von Störvariablen vorhanden sind, was für die Geltung von Hypothese 2 spricht.

#### **7.4.2 Vorhersage der Feedbackwirkung durch verschiedene Einflussfaktoren**

Wie in Abschnitt 7.3 dargestellt, wurden für folgende Variablen positive Effekte des Feedbacks gefunden: „Language Modeling“, „Geschlossene Fragen mit Antwortmöglichkeit“, „Ja/Nein-Fragen mit Antwortmöglichkeit“, „Offene Fragen mit Antwortmöglichkeit“, „Offene Fragen ohne Antwortmöglichkeit“, „Erzählaufforderungen mit Antwortmöglichkeit“, „Wiederholung ohne Korrektiv“, „Korrektive Wiederholung“, „Korrektive Erweiterung“, „Erweiterung ohne Korrektiv“ sowie „Anknüpfen“. In einem letzten Schritt wurde untersucht, ob die Veränderungen innerhalb der Feedbackgruppe von bestimmten personellen oder situativen Faktoren beeinflusst werden. In den folgenden Abschnitten werden die Ergebnisse der hierarchischen Regressionsanalysen beschrieben, bei denen immer im ersten Modell die jeweiligen Störvariablen und im zweiten Modell zusätzlich die interessierenden Einflussvariablen berücksichtigt wurden.

Zur Identifizierung potenzieller Störvariablen, die in den Regressionsanalysen kontrolliert werden sollten, wurden Korrelationen zwischen den Differenzwerten der sprachlichen Anregung (Kriterium) und verschiedenen persönlichen und situativen Merkmalen berechnet (siehe Tabelle 46, Anhang C8). Welche Störvariablen aufgrund eines signifikanten Zusam-

menhangs mit dem Differenzwert in die einzelnen Analysen einbezogen wurden, kann in den Ergebnistabellen in Anhang C9 eingesehen werden. In einigen Fällen stand die Leistung zu T1 in einem signifikanten Zusammenhang zu der Veränderung zwischen T1 und T2. Fachkräfte mit einer höheren Leistung zu T1 zeigen demnach erwartungsgemäß häufig geringere Verbesserungen. Vereinzelt zeigten sich auch signifikante Korrelationen mit der Funktion der Fachkraft sowie deren beruflichen Abschluss. Die Ergebnisse deuten an, dass Fachkräfte mit einem höheren Abschluss und einer höheren Funktion in manchen Aspekten eher vom Feedback profitiert haben. Der Feedbackgeber hatte in der Variable „Wiederholung ohne Korrektur“ und beim „Anknüpfen“ einen Einfluss auf die Höhe der Differenzwerte. Insgesamt ist anzumerken, dass die Leistung zu T1 der stärkste Prädiktor war und die Regressionsgewichte der anderen Störvariablen unter Berücksichtigung der Anfangsleistung häufig keine statistische Bedeutsamkeit erreichten.

### Selbstwirksamkeit und Einstellung gegenüber Feedback

Um herauszufinden, welchen Einfluss die Einstellung der teilnehmenden Fachkräfte gegenüber Feedback auf dessen Effekte hat, wurden die vier Skalen „Utility“, „Accountability“, „Social Awareness“ und „Feedback Self-Efficacy“ aus der „Feedback Orientation Scale“ (FOS; Linderbaum & Levy, 2010) zusätzlich zu den jeweiligen Störvariablen als Prädiktoren in die Regressionsanalysen aufgenommen.

Aufgrund bisheriger Forschung war davon auszugehen, dass eine höhere Selbstwirksamkeit zu größeren Feedbackeffekten führt. Dieser Befund findet sich in der vorliegenden Studie nicht. Die Variable „Feedback Self-Efficacy“ war unter Berücksichtigung der Störvariablen für keine der sprachlichen Verhaltensaspekte ein signifikanter Prädiktor (siehe Tabelle 47 in Anhang C9). Hypothese 3 kann daher nicht bestätigt werden.

Die Differenz in den offenen Fragen ohne Antwortmöglichkeit wird signifikant durch die Variable „Utility“ vorhergesagt ( $B = -3.79$ ,  $p = 0.01$ ). Insgesamt konnten die Variablen der Einstellung zu Feedback gegenüber Modell 1 (Gruppenleitung als Prädiktor) keinen zusätzlichen Beitrag zur Varianzaufklärung der Differenz in den offenen Fragen ohne Antwortmöglichkeit leisten ( $\Delta R^2 = 0.16$ ,  $p = 0.14$ ).

Die Variable „Social Awareness“ ist ein signifikant positiver Prädiktor für die Veränderung im „Language Modeling“ ( $B = 0.51$ ,  $p = 0.04^9$ ), wobei auch hier insgesamt keine statistisch bedeutsame Veränderung in der Varianzaufklärung durch die in Modell 2<sup>10</sup> einbezogenen Skalen der Einstellung zu Feedback zu verzeichnen ist ( $\Delta R^2 = 0.11$ ,  $p = 0.17$ ). In allen anderen Variablen der sprachlichen Anregung konnte keine signifikante Vorhersage durch die

---

<sup>9</sup> Aufgrund fehlender Homoskedastizität erfolgte eine Korrektur mithilfe des Bootstrap-Verfahrens.

<sup>10</sup> Als Störvariablen wurden die Leistung zu T1 sowie der Abschluss der Fachkraft an einer Fachschule für Kinderpflege einbezogen.

Skalen zur Einstellung gegenüber Feedback gefunden werden (siehe Tabelle 47 in Anhang C9). Die Einstellung gegenüber Feedback hatte somit keinen signifikanten Einfluss auf die Feedbackeffekte (Fragestellung 2).

### Zielorientierung

Basierend auf Ergebnissen von Studien aus der Feedback-Forschung wurde angenommen, dass Personen mit einer höheren Lernorientierung einen größeren Nutzen aus Feedback ziehen können und die Leistungsorientierung der Fachkräfte nicht mit den Feedbackeffekten zusammenhängt. Zur Überprüfung dieser Annahmen wurde die Einschätzung der Fachkräfte hinsichtlich ihrer „mastery approach goals“, „mastery avoidance goals“, „performance approach goals“ sowie „performance avoidance goals“ jeweils im zweiten Hierarchieschritt (zusätzlich zu den Störvariablen) berücksichtigt. Die Ergebnisse der Regressionsanalysen weisen darauf hin, dass die Veränderungen in den einzelnen Variablen der sprachlichen Anregung nicht durch die Ausprägung der Zielorientierung aufgeklärt werden können (siehe Tabelle 48 in Anhang C9). Für die Auswirkungen des Feedbackgesprächs spielt es dementsprechend keine Rolle, ob die teilnehmenden Fachkräfte eher lernorientiert oder eher leistungsorientiert waren. Die in dieser Studie erhobenen Daten können daher Hypothese 5, nicht jedoch Hypothese 4 stützen.

### Berufsbezogene Motivation

Ebenfalls keine zusätzliche signifikante Varianzaufklärung konnte für die berufsbezogene Motivation der Fachkräfte festgestellt werden (siehe Tabelle 49 in Anhang C9). In Bezug auf die dritte Fragestellung muss deshalb festgehalten werden, dass die allgemeine berufliche Motivation der Fachkräfte keinen Einfluss darauf hatte, wie viel diese sich nach dem Feedback in der sprachlichen Anregung verbessert haben.

### Reflexionsniveau

Hinsichtlich des Reflexionsniveaus der Fachkräfte ergeben die Regressionsanalysen nur für die offenen Fragen ohne Antwortmöglichkeit ein statistisch bedeutsames Regressionsgewicht ( $B = 2.36$ ,  $p = 0.03$ ). Ein höheres Reflexionsniveau der Fachkraft sagt demnach eine größere Steigerung in der Anzahl der gestellten offenen Fragen ohne Antwortmöglichkeit vorher. Gegenüber der Vorhersage durch die Position der Fachkraft als Gruppenleitung konnten durch das Reflexionsniveau weitere 10.6 % der Varianz aufgeklärt werden ( $p = 0.03$ ). Dies entspricht einem kleinen bis mittleren Effekt. Für die Varianzaufklärung der übrigen sprachlichen Verhaltensweisen spielt das Reflexionsniveau keine Rolle (siehe Tabelle 50 in Anhang C9). Insgesamt muss die vierte Fragestellung daher so beantwortet werden,

dass das durch den Feedbackgeber eingeschätzte Reflexionsniveau keine Wirkung auf die Effekte des Feedbacks zu haben scheint.

### Bewertung des Feedbacks durch die Fachkräfte

Die Bewertungen des Feedbacks hinsichtlich der Aspekte „Satisfaction“, „Utility“, „Knowledge“, „Application to practice“ und „Individual organizational results“ wurden ebenfalls als mögliche Prädiktoren in den Regressionsmodellen für die einzelnen Variablen der sprachlichen Anregung einbezogen (siehe Tabelle 51, Anhang C9).

Die Variable „Individual organizational results“ fungiert für die Differenzwerte in den offenen Fragen mit Antwortmöglichkeit ( $B = -3.10$ ,  $p = 0.04$ ) über den Einfluss der Position der Fachkraft<sup>11</sup> hinaus als signifikant negativer Prädiktor. Die Bewertung des Feedbacks insgesamt liefert keinen zusätzlichen signifikanten Beitrag zur Varianzaufklärung ( $\Delta R^2 = 0.12$ ,  $p = 0.42$ ). In einer nicht korrigierten Regression hängt die Differenz in der Anzahl an Erzählauforderungen mit Antwortmöglichkeit ( $B = -4.68$ ,  $p < 0.01$ ) auch mit der Variable „Individual organizational results“ zusammen<sup>12</sup>. Da für diese Variable keine Homoskedastizität vorliegt, wurde das Bootstrap-Verfahren angewendet. Dies hatte zur Folge, dass das Regressionsgewicht die statistische Bedeutsamkeit verfehlt ( $p = 0.09$ ). Gleiches gilt für die Regression der Differenzwerte in der Variable „Anknüpfen“ auf die Bewertung in Bezug auf das erworbene Wissen<sup>13</sup> (Variable „Knowledge“). Unter der Berücksichtigung der Variablen „Abschluss Fachhochschule“ und „Funktion als stellvertretende Leitung“ konnte für die Skala „Satisfaction“ in Bezug auf die Vorhersage der Veränderungen in den Erweiterungen ohne Korrektur ein signifikantes Regressionsgewicht festgestellt werden ( $B = -6.33$ ,  $p = 0.01$ ). Durch Hinzunahme der fünf Skalen zur Bewertung des Feedbacks konnten insgesamt zusätzlich 23.2 % der Varianz aufgeklärt werden, was jedoch mit einem  $p = 0.08$  nicht signifikant ist. Da nur sehr vereinzelt die Differenzwerte der sprachlichen Anregung durch die Bewertung des Feedbacks vorhergesagt wurden, kann insgesamt nicht von einem bedeutenden Zusammenhang mit den Effekten des Feedbacks ausgegangen werden.

### Technische Probleme

Da bei einigen Feedbackgesprächen technische Probleme aufgetreten sind, wurde untersucht, ob diese die Effekte des Feedbacks beeinflusst haben (Fragestellung 6). Da für keine Variable der sprachlichen Anregung eine signifikante Varianzaufklärung durch das Vorhandensein von technischen Schwierigkeiten ermittelt wurde, ist davon nicht auszugehen (siehe Tabelle 52, Anhang C9).

---

<sup>11</sup> Funktion als Leitung als Störvariable

<sup>12</sup> Funktion als stellv. Leitung als Störvariable

<sup>13</sup> Unter Berücksichtigung von Feedbackgeber und Funktion als stellv. Leitung als Störvariablen:  $B = -1.33$ , ohne Bootstrap-Korrektur:  $p = 0.02$ , mit Bootstrap-Korrektur:  $p = 0.06$



## 8 Diskussion der Ergebnisse und Ableitung von Implikationen

Im Rahmen des Projekts „BiSS-Evaluation“ wurde von der Autorin ein Konzept entwickelt, mit dem man frühpädagogischen Fachkräften Feedback zu ihrem sprachlichen Interaktionsverhalten geben kann. Inhaltliche Grundlage dafür bilden Videosequenzen aus dem Alltag der Fachkraft, die gemeinsam mit ihr reflektiert werden. Die Feedbackgespräche finden online mithilfe einer Videokonferenzsoftware statt. Anlass für die Konzeptionsentwicklung waren Befunde empirischer Studien zur Interaktionsqualität und deren Optimierung. Vor allem hinsichtlich lernunterstützender Interaktionen von Fachkräften deuten die Ergebnisse zahlreicher Studien auf Entwicklungsbedarf hin (z.B. Helmerhorst et al., 2017; Wirts et al., 2017). Das Qualitätsniveau der sprachlichen Lernunterstützung liegt im Durchschnitt in Deutschland (z.B. Wirts et al., 2017) im niedrigen bis mittleren Bereich. Die Ergebnisse von verschiedenen Meta-Analysen bestätigen die wichtige Rolle von Fort- und Weiterbildungen für die Qualitätsentwicklung (z.B. Egert, 2015; Markussen-Brown et al., 2017). In der klassischen Lehr-/Lernforschung wurden in vielen Studien positive Effekte von Feedback bestätigt (z.B. Hattie & Timperley, 2007). Für eine empirische Überprüfung des entwickelten Konzepts wurden in einem Prä-Post-Kontrollgruppendesign die Effekte eines Feedbackgesprächs auf das sprachliche Interaktionsverhalten von frühpädagogischen Fachkräften während einer Bilderbuchbetrachtung untersucht. Es zeigten sich für einen großen Teil der relevanten Variablen signifikant größere Verbesserungen in der Feedbackgruppe.

In den folgenden Abschnitten sollen die Erkenntnisse und Ergebnisse der vorliegenden Arbeit zusammengefasst und diskutiert werden. Im Rahmen dieser Diskussion werden die Befunde interpretiert, kritisch reflektiert und in die bisherige Literatur eingebettet. Darauf aufbauend werden Implikationen für die weitere Forschung und die Praxis abgeleitet.

### 8.1 Erprobung des Feedback-Konzepts

Vor der Diskussion der empirischen Ergebnisse soll das entstandene Feedbackkonzept noch einmal zusammenfassend beschrieben und die Erfahrungen mit dessen Umsetzung reflektiert werden.

#### Feedback-Konzept

Der Zweck eines Feedbacks kann im Vergleich von einem Ist-Zustand mit einem Soll-Zustand bestehen (Müller & Ditton, 2014). Das sprachliche Interaktionsverhalten der Fach-

kraft während einer Bilderbuchbetrachtung ist in diesem Fall der Ist-Stand. Im entwickelten Konzept wird dieses mit einem sachlichen Kriterium verglichen, nämlich den Indikatoren der Dimension „Language Modeling“ aus dem „Classroom Assessment Scoring System“ von Pianta und La Paro et al. (2008). Die Grundlage für das Feedback bildet demnach die vorherige Einschätzung des Ist-Stands. Dabei stellt der ermittelte „CLASS-Score“ für die Dimension „Language Modeling“ nur eine Orientierung für den Feedbackgeber dar. Das Feedback bezieht sich ausschließlich auf konkrete Verhaltensbeispiele für die Umsetzung der Indikatoren sprachlicher Anregung. Da die Ermittlung des Ist-Stands eine wesentliche Voraussetzung für Feedback ist, muss sichergestellt sein, dass der Feedbackgeber in der Lage ist, das Verhalten der Fachkräfte objektiv und reliabel einzuschätzen. Nur so kann eine vergleichbare Qualität des Feedbacks gewährleistet werden.

Der Ablauf des Feedbackgesprächs lehnt sich an das Modell von Hattie und Timperley (2007) an. Nach Ansicht der Autoren sollten im Rahmen eines Feedbacks in erster Linie drei Fragen beantwortet werden: „Where am I going?“, „How am I going?“ und „Where to next?“. Die Beantwortung dieser Fragen steht daher im Fokus des Feedbacks und entspricht den ersten drei Bausteinen in der Durchführung. Da Feedbackinterventionen mit Zielsetzungsmaßnahmen höhere Effekte aufweisen als Formate ohne solche Zielformulierungen (Hattie, 1999), besteht der vierte Baustein des Konzepts aus einer Zielvereinbarung. Die einzelnen Bausteine werden im Folgenden in Kurzform dargestellt und anschließend anhand der in dieser Studie gewonnenen Erkenntnisse diskutiert.

Der erste Baustein dient der Klärung des Soll-Zustands („Where am I going?“). Dafür wurden der Fachkraft zunächst Kriterien für eine erfolgreiche sprachliche Anregung mithilfe einer ins Deutsche übersetzten Übersicht über die CLASS-Indikatoren (Häufige Gespräche, Offene Fragen, Wiederholen und Erweitern, Handlungsbegleitendes Sprechen sowie Differenzierte Sprache) vorgestellt.

Der Hauptteil des Feedbacks widmet sich der Beantwortung der anderen beiden Fragen. Die Frage „How am I going“ zielt auf die Rückmeldung bzw. Reflexion des Ist-Stands ab und entspricht daher am ehesten einem Feedback im eigentlich Wortsinn. Mit dem Ziel, eine aktive Konstruktion von Wissen anzuregen, wird gemeinsam mit der Fachkraft reflektiert, an welchen Stellen sie das angestrebte Verhalten bereits sehr gut umsetzt und wo es noch Optimierungspotenzial gibt. Dafür werden zunächst gelungene Beispiele angeschaut, um der Fachkraft eine Vorstellung vom Indikator und einer möglichen Umsetzung zu vermitteln und zu zeigen, dass sie prinzipiell in der Lage dazu ist, die jeweilige sprachförderliche Strategie einzusetzen. Anschließend wird anhand von Situationen mit Optimierungspotenzial reflektiert, wie das entsprechende Verhalten noch besser oder noch häufiger umgesetzt werden könnte. Dabei wird die Fachkraft mithilfe von Fragen, welche beispielhaft im Leitfa-

den formuliert sind, zur Reflexion angeregt (z.B. „Vergleichen Sie die Reaktion der Kinder auf diese beiden Fragen. Was fällt Ihnen auf?“). Sobald Optimierungspotenzial identifiziert wird, wird direkt daran anknüpfend erarbeitet, wie eine bessere Umsetzung des entsprechenden Verhaltens gelingen könnte. Auch an dieser Stelle wird mit Reflexionsfragen gearbeitet (z.B. „Wie könnte man diese geschlossene Frage offen formulieren?“). Als Ergänzung oder Hilfestellung können bei beiden Schritten Hinweise oder Beispiele gegeben werden. Um den Fachkräften eine möglichst gute Verknüpfung der Feedbackinhalte mit dem eigenen Verhalten zu ermöglichen, wurde das Feedback-Konzept videobasiert konzipiert. Im zweiten und dritten Baustein („How am I going?“ und „Where to next?“) wird daher mit ausgewählten Sequenzen einer vorher angefertigten und ausgewerteten Videoaufnahme der pädagogischen Fachkraft gearbeitet.

In einem letzten Schritt werden mit der Fachkraft aufbauend auf dem Feedbackgespräch Ziele festgelegt. Die Ziele werden dabei von der Fachkraft mit Unterstützung des Feedbackgebers formuliert, vom Feedbackgeber schriftlich festgehalten und anschließend an die Fachkraft übermittelt.

Bezugnehmend auf die Klassifizierung von Narciss (2006) vereint das Feedback somit verschiedene Feedbackformen miteinander (siehe Tabelle 1, S. 40). Zusätzlich zu einfachen Feedbackinhalten, wie beispielsweise der Erfolg bei der Umsetzung des entsprechenden Verhaltens („Knowledge of result“), werden im Rahmen des Feedbacks vor allem auch stärker elaborierte Informationen geliefert bzw. erarbeitet (z.B. Hinweise auf die Anforderungen, Erklärung von Begriffen, Diskussion der gemachten „Fehler“, Lösungshinweise).

### Ablauf der Feedbackgespräche

Mit der beschriebenen Vorgehensweise wurden in den für diese Studie durchgeführten Feedbackgesprächen (maximal) drei der fünf CLASS-Indikatoren reflektiert.

Die inhaltliche Gliederung des Konzepts konnte in den Feedbackgesprächen gut umgesetzt werden. Die Durchführung mit der Arbeitsvorlage hat sich als sehr praktikabel erwiesen. Die Strukturierung des Feedbacks wurde auch von den teilnehmenden Fachkräften gut angenommen. Laut einer zu T2 abgegebenen schriftlichen Einschätzung hat den Fachkräften am besten die Analyse der Videosequenzen und die damit verbundene Möglichkeit, das eigene Verhalten „von außen“ zu betrachten und gemeinsam zu reflektieren, gefallen. Aber auch die Darstellung der Kriterien für ein sprachförderliches Verhalten zu Beginn sowie die schriftliche Formulierung von Zielen am Ende des Feedbackgesprächs wurden von einigen Fachkräften explizit als positive und hilfreiche Aspekte genannt.

Um den Fachkräften einen Überblick über die wichtigsten Aspekte von sprachlicher Anregung im Kita-Alltag zu geben, wurden im Rahmen des ersten Bausteins („Where am I going?“) auch Strategien thematisiert, die für die Situation der Bilderbuchbetrachtung nicht

primär von Interesse sind (z.B. Handlungsbegleitendes Sprechen). Da dies zum Teil zu Irritationen geführt hat, wäre für zukünftige Anwendungen auch denkbar, immer nur Aspekte zu besprechen, die für die jeweilige Alltagssituation relevant sind.

Im Sinne der positiven Verstärkung liegt der Fokus bei der Bearbeitung des ersten Indikators auf dem, was bereits gut umgesetzt wurde. Hier wurde am häufigsten das Wiederholen und Erweitern kindlicher Äußerungen thematisiert, was darauf hinweist, dass dieser Aspekt von den meisten Fachkräften bereits gut umgesetzt wird. Als wichtig zu erwähnen, erscheint an dieser Stelle, dass bei ca. einem Fünftel der Fachkräfte nicht genügend positive Beispiele für sprachliche Anregung vorhanden waren und daraufhin ein nicht sprachlicher Aspekt (z.B. Atmosphäre in der Gruppe) zur positiven Verstärkung gewählt wurde. Obwohl bei Bilderbuchbetrachtungen in einigen Studien ein höheres Niveau sprachlicher Anregung beobachtet wurde als in anderen Alltagssituationen (z.B. Massey et al., 2008), scheint es angesichts dieses Ergebnisses dennoch Bedarf für Qualitätsentwicklung zu geben. Als Indikator 2 und 3 wurden vor allem die Gespräche mit den Kindern und das Stellen von offenen Fragen besprochen. Basierend auf positiven Beispielen der Umsetzung wurde bei diesen Indikatoren auch das Optimierungspotenzial in den Blick genommen. Damit hebt sich das Vorgehen von anderen Formaten ab, die im Sinne der Verstärkung vorrangig positive Verhaltensbeispiele reflektieren (z.B. „Video Interaction Guidance“, Fukkink & Tavecchio, 2010). Der hier gewählte Ansatz hat insofern Bestätigung gefunden, als von vielen Fachkräften positiv bewertet wurde, dass sowohl Stärken als auch Schwächen diskutiert wurden. Wie bereits erwähnt, hat den Fachkräften die Reflexion der Indikatoren mit den Videosequenzen am meisten gefallen. Hierbei muss berücksichtigt werden, dass es sich aufgrund der Freiwilligkeit der Teilnahme um eine selektive Stichprobe handelt. Es kann davon ausgegangen werden, dass die teilnehmenden Fachkräfte prinzipiell eine offene Einstellung gegenüber der Arbeit mit Videomaterial haben. Einige Fachkräfte haben die Beteiligung an dem Projekt jedoch gerade wegen der Videoaufnahmen abgelehnt. Während der Stichprobenakquise zeigten sich zum Teil große Vorbehalte von Fachkräften und Eltern gegenüber der Arbeit mit Videoaufnahmen. Eine umfangreiche Aufklärungsarbeit (z.B. zum Thema Datenschutz) könnte dafür sorgen, dass manche Ängste und Sorgen abgebaut bzw. gar nicht erst aufgebaut werden.

Im Zuge der Zielvereinbarung wurde mit Abstand am häufigsten das Stellen von offenen Fragen als Ziel formuliert. Viele Fachkräfte haben sich zudem vorgenommen, mehr mit den Kindern in den Dialog gehen, auf die Äußerungen der Kinder einzugehen bzw. daran anzuknüpfen sowie weniger vorwegzunehmen und mehr die Kinder sprechen zu lassen. Die formulierten Ziele setzen somit ausschließlich auf der Handlungsebene an (siehe Kapitel 5.2.2). In der „goal setting theory“ von Locke und Latham (1990) wird betont, dass spezifi-

sche, herausfordernde Ziele lernförderlicher sind als allgemein gehaltene „do your best“-Ziele (siehe Kapitel 3.2.4). Da die meisten Fachkräfte am Ende der durchgeführten Feedbackgespräche selbständig sehr konkrete und dem Inhalt des Gesprächs entsprechende Ziele formuliert haben, kann die Zielvereinbarung hinsichtlich dieses Aspekts insgesamt als gelungen betrachtet werden. Die in der Literatur betonte persönliche Bedeutung und Akzeptanz der Ziele wurde dadurch sichergestellt, dass die Fachkräfte selbst ihre Ziele formuliert haben. Da die Zielvereinbarung nicht als Untersuchungsbedingung manipuliert wurde, kann nicht beantwortet werden, ob sie einen zusätzlichen lernförderlichen Effekt hatte. Dies wäre somit eine interessante Fragestellung für künftige Forschungsvorhaben.

Die Gespräche haben im Durchschnitt ca. eine Stunde gedauert. Die Erfahrung mit der Durchführung der Feedbackgespräche in dieser Studie hat gezeigt, dass dies eine angemessene Zeitspanne ist, da sich Gespräche von einer Stunde noch gut in den Alltag der Fachkräfte integrieren lassen. Zu diesem Schluss sind auch Werner et al. (2018) in ihrer Untersuchung gekommen, in der sie die Effekte eines Trainings mit Videofeedback zur Feinfühligkeit von pädagogischen Fachkräften geprüft haben. Für die Dauer von einer Stunde sind die Fachkräfte in der Regel in der Gruppe verzichtbar oder können durch KollegInnen vertreten werden. Alternativ könnte das Feedback auch in der Mittagspause stattfinden. Auch die Videoaufnahmen konnten in den meisten Fällen ohne großen zusätzlichen Aufwand angefertigt werden. Im Rahmen dieser Intervention wurden die Bilderbuchbetrachtungen vom jeweiligen Feedbackgeber videografiert. Dahinter steckte die Absicht, vor dem Feedback einen persönlichen Kontakt zur Fachkraft herzustellen. Um die damit verbundenen zeitlichen Ressourcen einzusparen, wäre es auch denkbar, dass die Fachkräfte selbst oder eine Kollegin die Aufnahme machen – vor allem, wenn bereits ein persönlicher Kontakt besteht. Bei „MyTeachingPartner“ filmen sich die Fachkräfte beispielsweise selbst und schicken die Aufnahme anschließend über ein Online-Portal an ihren Trainer (Pianta, Mashburn et al., 2008).

Für das Feedback-Konzept wurde bewusst die mündliche Kommunikationsform gewählt, weil ein aktiver Einbezug der Fachkräfte und ein gemeinsamer Austausch möglich gemacht werden sollte. In der Meta-Analyse von Kluger und DeNisi (1996) hatten allerdings Studien, bei denen das Feedback mündlich stattfand, signifikant kleinere Effekte als schriftliche Feedbackformen. Eine mögliche Ursache dafür könnte sein, dass bei einem mündlichen Feedback personale Faktoren des Feedbackgebers sowie deren Zusammenspiel mit Merkmalen des Feedbackempfängers stärker zum Tragen kommen. Die Überprüfung dieser Hypothese war nicht Gegenstand dieser Arbeit, weshalb keine empirisch belastbaren Aussagen dazu getroffen werden können. Die Reaktionen und Beurteilungen der Fachkräfte sprechen jedoch dafür, dass vor allem die gemeinsame Reflexion des Verhaltens als

wichtiger und nützlicher Bestandteil des Feedbacks empfunden wurde. Da dies bei einem schriftlichen Feedback nicht möglich ist, scheint die mündliche Kommunikationsform für diese Zielgruppe erfolgsversprechender. Untersuchungen zum Vergleich von schriftlichen und mündlichen Feedbackformen bei frühpädagogischen Fachkräften könnten dazu weiter Aufschluss geben. Besonders wichtig für ein mündliches Feedbackgespräch scheint zu sein, wie das Feedback übermittelt wird. Ein großer Teil der Fachkräfte hat die angenehme Atmosphäre sowie die freundliche und wertschätzende Haltung und Kommunikation durch den Feedbackgeber als besonders positiven Aspekt des Feedbackgesprächs genannt. Während die Bedeutung von Kommunikation und Gesprächsführung im Kontext von Coaching und im therapeutischen Setting bereits allgegenwärtig ist, wird dies in der vorhandenen Literatur zu Feedback und dessen Effekten bisher kaum thematisiert.

### Technische Umsetzung

Da persönliche Gespräche vor Ort große finanzielle und zeitliche Ressourcen beanspruchen, fanden die Feedbackgespräche online mithilfe einer Videokonferenzsoftware statt.

Die technische Umsetzung dieser Vorgehensweise hat sich bereits in der Entwicklungsphase als eine Herausforderung herausgestellt. Wie in Kapitel 5.1 beschrieben, muss eine mögliche Software strengen Datenschutzauflagen genügen sowie die Möglichkeit bieten, während eines Gesprächs mit Webcam-Übertragung Videosequenzen gemeinsam anzuschauen. Von den zahlreichen getesteten Software-Lösungen hat nur die kostenpflichtige E-Learning-Plattform vitero die erforderlichen Kriterien erfüllt. Bei den meisten anderen Angeboten war entweder der Datenschutz nicht sichergestellt oder eine Tonübertragung während des Abspielens der Videosequenzen war nicht möglich. Kostenfreie Software mit den entsprechenden technischen Möglichkeiten würde die Umsetzbarkeit von online-basiertem Videofeedback im Feld enorm erleichtern.

In der vorliegenden Studie konnten alle Feedbackgespräche durchgeführt werden. Prinzipiell hat die technische Umsetzung demnach gut funktioniert. Allerdings sind während der Feedbackgespräche relativ häufig technische Probleme aufgetreten (siehe Kapitel 7.1). Bei einigen Fachkräften war die Ton- bzw. Webcam-Übertragung verzögert und zum Teil musste die Sitzung zwischendurch unterbrochen und neu gestartet werden. Es ist anzunehmen, dass eine schlechte Internetverbindung in den Kindertageseinrichtungen verantwortlich für diese Schwierigkeiten war. Kein Hinderungsgrund, aber dennoch relevant für die Umsetzung des Feedbacks war der Umstand, dass viele Kitas nur über einen Computer im Büro der Leitung verfügen. Dies hat vor allem die zeitliche Planung der Feedbackgespräche erschwert. Häufig kam es auch zu Problemen aufgrund der stark eingeschränkten Nutzungsrechte am Computer. In der Regel konnten diese mithilfe der jeweiligen IT-Abteilung gelöst werden. In zwei Fällen mussten die Fachkräfte allerdings auf ihren privaten Computer zu

Hause ausweichen. Wenn Feedbackgespräche dieser Art in Zukunft eine Fortbildungsmöglichkeit darstellen sollen, muss der Zugang zu einem gut funktionierenden Computer mit stabiler Internetverbindung sichergestellt werden.

Die Bedeutung der technischen Umsetzung findet sich zum Teil auch in den Bewertungen der Fachkräfte wieder. Einige Fachkräfte haben zurück gemeldet, dass Online-Feedback eine gute Möglichkeit der Weiterbildung ist, sofern die Technik funktioniert. Trotz der genannten Probleme wurde das online-basierte Feedback insgesamt sehr positiv wahrgenommen. Die Vorteile wurden vor allem in der Zeitersparnis, der Effektivität sowie in der zeitlichen Flexibilität gesehen.

Zusammenfassend kann festgehalten werden, dass sich das Feedback-Konzept trotz technischer Probleme in der praktischen Umsetzung bewährt hat. Auch die Fachkräfte haben das Feedbackgespräch insgesamt sehr gut bewertet. Die schriftliche Befragung mit dem „Questionnaire for Professional Training Evaluation“ (Grohmann & Kauffeld, 2013) ergab eine positive Einschätzung der Fachkräfte hinsichtlich ihrer Zufriedenheit, dem Nutzen für ihre pädagogische Arbeit, dem erworbenen Wissen und dessen praktischer Anwendung sowie der Relevanz für ihre allgemeine Arbeit in der Kita. Betrachtet man die hohe Zufriedenheit der Teilnehmer als ein wesentliches Kriterium für den Erfolg von Bildungsmaßnahmen, kann das durchgeführte Online-Feedback als sinnvolles Fortbildungsformat betrachtet werden. Die Ergebnisse in Bezug auf die beobachtbaren Auswirkungen des Feedbacks auf Performanz-Ebene werden in den nächsten Abschnitten diskutiert.

## 8.2 Effekte des Feedbackgesprächs

Das Ziel der Arbeit bestand vor allem darin, die Auswirkungen des Feedbacks auf das sprachliche Interaktionsverhalten der Fachkräfte zu untersuchen.

### „Language Modeling“

Zur Bewertung der sprachlichen Anregung wurde die Dimension „Language Modeling“ aus dem „Classroom Assessment Scoring System“ von Pianta und La Paro et al. (2008) herangezogen. Bei den Doppelkodierungen konnte eine zufriedenstellende Übereinstimmung zwischen den Ratern erreicht werden. Die sprachliche Anregung durch die Fachkräfte zu T1 lag durchschnittlich im unteren bis mittleren Bereich der sieben-stufigen Skala. Dies entspricht dem Niveau in anderen deutschen (Kammermeyer et al., 2013) und internationalen Studien (z.B. Justice et al., 2008). Der Entwicklungsbedarf bezüglich der sprachlichen Anregung im Kindergarten wurde daher auch mit den Daten dieser Studie bestätigt. In Bezug auf die gezeigte sprachliche Anregung zu T1 bestand kein bedeutender Unterschied zwischen Kontroll- und Feedbackgruppe. Was die Veränderungen zu T2 angeht, deuten die

Ergebnisse der zweifaktoriellen Varianzanalyse mit Messwiederholung auf eine signifikante Interaktion zwischen Messzeitpunkt und Bedingung mit großer Effektstärke hin. Dies bedeutet, dass sich die beiden Gruppen in ihren Differenzwerten zwischen T1 und T2 relativ stark voneinander unterscheiden. Die deskriptiven Kennwerte zeigen diesbezüglich, dass sich die Kontrollgruppe von T1 zu T2 leicht verschlechtert. Die Auswertung des *t*-Tests für abhängige Stichproben ergab jedoch keine statistische Bedeutsamkeit für diesen Unterschied. Die Fachkräfte in der Feedbackgruppe hatten zu T2 hingegen einen höheren Mittelwert als zu T1. Im Durchschnitt haben sie sich um ca. 0.75 Punkte auf der sieben-stufigen Skala verbessert, was einer Effektstärke von  $g = 0.78$  und somit einem großen Effekt entspricht.

Sowohl die Meta-Analysen zur Wirksamkeit von Fortbildungen für frühpädagogische Fachkräfte (z.B. Fukkink & Lont, 2007) als auch die Meta-Analysen zu den Effekten von Feedback (z.B. Kluger & DeNisi, 1996) stellen durchschnittlich mittelgroße Effekte fest. Dabei hat ein nicht zu vernachlässigender Teil sogar negative Effekte. Das durchgeführte online-basierte Videofeedback kann demnach größere Effekte aufweisen als der Durchschnitt an Fortbildungen und Feedback-Interventionen. Für CLASS-basierte Trainings berichtet Egert (2015) hingegen große Effekte. Auch Fortbildungen mit Videofeedback oder Coaching zeigten im Mittel große Effekte (Egert, 2015). In Meta-Analysen zu Feedbackeffekten wurden ebenfalls bei videogestützten Verfahren höhere Effekte festgestellt als bei Methoden ohne Videoanalyse (Hattie, 1999). Diesbezüglich sind die Effektstärken vergleichbar. Allerdings muss dabei berücksichtigt werden, dass es sich bei den in den Meta-Analysen einbezogenen Studien um Formate mit einem deutlich größeren zeitlichen Umfang handelt. Bei Egert (2015) dauerte die kürzeste Intervention insgesamt vier Stunden, bei Markussen-Brown et al. (2017) waren es sechs Stunden. In beiden Fällen erstreckten sich alle Fortbildungen über mehrere Tage. Einige waren auch auf mehrere Wochen oder Monate angelegt. Pianta und Mashburn et al. (2008) erreichten mit ihrer mehrmaligen videogestützten Beratung sogar nur kleine Effekte im „Language Modeling“. Unter Berücksichtigung der Tatsache, dass das Feedback in der vorliegenden Studie eine einmalige Intervention mit einem Zeitumfang von ca. einer Stunde war, sind die beobachteten großen Effekte als umso bedeutender einzuschätzen.

In der Meta-Analyse von Egert (2015) zeichneten sich besonders effektive Formate unter anderem dadurch aus, dass sie Praxisaufgaben als Arbeitsaufträge enthielten (z.B. das Üben von speziellen Sprachfördertechniken). Da in der vorliegenden Studie ausschließlich die Effekte des Feedbackgesprächs untersucht werden sollten, wurde auf eine entsprechende Anweisung verzichtet. Mit Blick auf die Ergebnisse der Meta-Analyse wäre es jedoch denkbar, dass die Effekte des Feedbacks durch zusätzliche Praxisaufgaben noch ge-



steigert oder zumindest verfestigt werden könnten. Dies zu überprüfen, ist Aufgabe künftiger Untersuchungen.

Um direkte Veränderungen auf der Verhaltensebene abzubilden, ist die Bewertung der Dimension „Language Modeling“ zu global (Mayer & Beckh, 2018). Die Effekte des Feedbacks wurden daher auch in Bezug auf die konkrete Anzahl sprachanregender Verhaltensweisen überprüft.

### Anzahl sprachlicher Verhaltensweisen

Das sprachliche Verhalten der Fachkräfte wurde mit einem Kodier-System erfasst, welches sich an die Indikatoren der Dimension „Language Modeling“ anlehnt und diese weiter ausdifferenziert. Mithilfe des Schemas wurde die Anzahl an sprachlichen Verhaltensweisen während der Bilderbuchbetrachtung gezählt.

#### *Auswertung mit dem Kodier-Schema*

Angelehnt an die Operationalisierung von Siraj-Blatchford und Manni (2008) erfolgte eine Kategorisierung in die folgenden Fragetypen: geschlossene Fragen, Ja/Nein-Fragen, vermeintliche Fragen und offene Fragen. Erzählaufforderungen wurden ergänzend dazu aufgenommen. Dabei wurde immer in „mit Antwortmöglichkeit“ und „ohne Antwortmöglichkeit“ für die Kinder unterschieden. Entsprechend der Indikatoren in der CLASS wurden darüber hinaus das handlungsbegleitende Sprechen, das Wiederholen und Erweitern kindlicher Äußerungen sowie längere Gespräche und das Anknüpfen an Äußerungen der Kinder im Kategoriensystem berücksichtigt. In Bezug auf das Wiederholen und Erweitern wurde zusätzlich danach differenziert, ob es sich um eine korrektive oder nicht korrektive Form handelt.

Im Rahmen der Doppelkodierungen konnte bei den meisten Variablen eine ausreichend hohe Inter-Rater-Reliabilität festgestellt werden. Für einzelne Variablen ergaben die Analysen eine sehr geringe Reliabilität bzw. eine ICC von 0. Bei den längeren Gesprächen, den Erzählaufforderungen ohne Antwortmöglichkeit und dem handlungsbegleitenden Sprechen könnte dies unter anderem mit dem sehr seltenen Auftreten dieser Verhaltensweisen zu tun haben. Wenn die Varianz einer Variable sehr gering ist, fällt es umso schwerer diese Varianz sauber aufzuklären (Wirtz & Caspar, 2002). Die reliable Auswertung der Videoaufnahmen hinsichtlich der einzelnen Verhaltensweisen stellte allgemein eine große Herausforderung dar. Trotz einer sehr genauen Beschreibung aller Codes bleibt immer ein gewisser Interpretationsspielraum. Dies ist zum Beispiel bei der Frage nach der Antwortmöglichkeit aufgefallen. So kann die Entscheidung, ob einem Kind ausreichend Gelegenheit für eine Antwort gegeben wurde, individuell sehr unterschiedlich ausfallen und hängt zudem auch vom jeweiligen Kind ab. Da die Videoaufnahmen direkt ausgewertet wurden, ergaben

sich zudem Schwierigkeiten mit dem akustischen Verständnis. Wenn beispielsweise nicht klar ist, ob die Äußerung vom Kind eine Aussage oder eine Frage war, könnte die Reaktion der Fachkraft als Wiederholung bzw. Erweiterung oder als Anknüpfen gedeutet werden. Gleiches gilt für die „Fehlerhaftigkeit“ einer kindlichen Äußerung und die daraus resultierende Frage nach korrektiv oder nicht korrektiv. Mithilfe von Transkripten könnte dieser Schwierigkeit begegnet werden. Dies war in der vorliegenden Studie aufgrund der Vielzahl an Videoaufnahmen nicht möglich. Darüber hinaus wäre zum Beispiel die Einschätzung der Antwortmöglichkeit mit einem Transkript nur schwer möglich. Ein weiteres Problem stellten umgangssprachliche bzw. dialektale Besonderheiten dar. In der vorliegenden Studie haben die Fachkräfte zum Beispiel häufig eine Aussage durch das Wort „gell“ ergänzt. Wenn diese eindeutig als Frage intoniert war und die Antwort der Kinder abgewartet wurde, wurde sie als „Ja/Nein-Frage“ gezählt. Bei manchen Fachkräften wurde das „gell“ jedoch in fast jedem Satz verwendet. In diesen Fällen war es sehr schwierig zu entscheiden, ob es in dem jeweiligen Fall wirklich als Frage gemeint war oder eher als Füllwort diente. Gerade wegen der genannten Herausforderungen wird die erreichte Reliabilität als zufriedenstellend bewertet. Es hat sich gezeigt, dass dafür eine sehr detaillierte Beschreibung mit vielen Beispielen sowie eine intensive Auseinandersetzung mit den Codes notwendig ist. Gleiches gilt für die Einschätzung mit der CLASS. Da ein Feedback immer die Erfassung des Ist-Stands mit eben solchen Variablen oder Instrumenten voraussetzt, sollte bei der Qualifizierung von Feedbackgebern unbedingt eine gute Schulung hinsichtlich der relevanten Merkmale und deren Beurteilung erfolgen. Unterstützt wird diese Forderung durch bestehende Befunde aus der Forschung. In der Meta-Analyse von Egert (2015) stellte sich heraus, dass besonders effektive Trainings gut ausgebildete und speziell geschulte Trainer eingesetzt haben. Auch in der Studie von Feys, Anseel und Wille (2011) haben die Teilnehmer dem Feedback mehr Bedeutung beigemessen, wenn sie wussten, dass es von geschulten Beobachtern übermittelt wurde.

### *Deskriptive Ergebnisse*

Betrachtet man die gestellten Fragen zu T1, fällt auf, dass mit Abstand am häufigsten Ja/Nein-Fragen mit Antwortmöglichkeit gestellt wurden. Geschlossene und offene Fragen mit Antwortmöglichkeit kamen in etwa gleich häufig vor. Von insgesamt 5788 gestellten Fragen zu T1 waren etwas mehr als ein Drittel Ja/Nein-Fragen mit Antwortmöglichkeit und jeweils ca. ein Fünftel geschlossene Fragen mit Antwortmöglichkeit und offene Fragen mit Antwortmöglichkeit. Siraj-Blatchford und Manni (2008) konnten von 5808 ausgewerteten Fragen aus dem Alltag von pädagogischen Fachkräften 45.2 % als geschlossen und nur 5.5 % als offen identifizieren. Ja/Nein-Fragen mit Antwortmöglichkeit haben 11.29 % der Fragen ausgemacht. Verglichen damit ist der Anteil an offenen Fragen in der vorliegenden

Studie demnach deutlich höher. Dies könnte daran liegen, dass in den Daten von Siraj-Blatchford und Manni (2008) jeweils der gesamte Vormittag berücksichtigt wurde und eventuell während einer Bilderbuchbetrachtung allgemein mehr offene Fragen gestellt werden als in anderen Alltagssituationen. Dafür spricht, dass bei Milburn et al. (2014) und Simon und Sachse (2011) eine mit den vorliegenden Daten ähnliche Anzahl an offenen Fragen während einer Bilderbuchbetrachtung beobachtet wurde. Diese Annahme bestätigend, konnten Massey et al. (2008) zeigen, dass in Bilderbuchsituationen mehr offene Fragen gestellt wurden als in anderen Situationen. Hierbei muss bedacht werden, dass sehr unterschiedliche Definitionen von offenen und geschlossenen Fragen in den Untersuchungen vorliegen. Dies macht den Vergleich von Daten verschiedener Studien zum Teil schwierig, weshalb er mit Vorsicht zu interpretieren ist. Erzählaufforderungen kamen deutlich seltener vor als Fragen und spielen scheinbar in den sprachlichen Strategien der Fachkräfte nur eine nebensächliche Rolle. Hinsichtlich der Fragen und Erzählaufforderungen wurde bei der Auswertung in „mit Antwortmöglichkeit“ und „ohne Antwortmöglichkeit“ (z.B. zwei aufeinander folgende Fragen oder Fachkraft gibt selbst Antwort) unterschieden, da nur erstere tatsächlich einen sprachanregenden Charakter haben. Fragen oder Erzählaufforderungen ohne Antwortmöglichkeit waren insgesamt viel weniger häufig vertreten als die mit Antwortmöglichkeit. Die gewonnenen Daten deuten darauf hin, dass die Kinder in der Regel ausreichend Möglichkeit zum Antworten bekommen, aber dennoch regelmäßig auch Fragen oder Aufforderungen ohne Antwortmöglichkeit gestellt werden. Dieser Aspekt war bisher nur selten Gegenstand von wissenschaftlichen Untersuchungen. Häufig wird die Antwortmöglichkeit gar nicht berücksichtigt (de Rivera et al., 2005) oder Fragen ohne Antwortmöglichkeit werden bewusst nicht mit gezählt (Siraj-Blatchford & Manni, 2008). Die vorliegenden Daten können einen ersten Einblick zu diesem Thema geben, es bedarf jedoch weiterer Studien zur Replikation und Ergänzung der Ergebnisse, z.B. durch die Analyse in anderen Alltagssituationen.

Während der Bilderbuchbetrachtungen fanden fast keine längeren Gespräche (mind. vier Sprecherwechsel) mit einzelnen Kindern statt. In der Studie von Milburn et al. (2014) hatten die Fachkräfte während einer Bilderbuchbetrachtung im Schnitt weit mehr als zehn Gespräche mit mindestens fünf Sprecherwechseln mit den Kindern. Allerdings wurden dabei alle Gespräche zwischen der Fachkraft und einem oder mehreren Kindern berücksichtigt. Längere Gespräche mit einer Kindergruppe scheinen demnach offensichtlich bedeutend öfter stattzufinden. Die Beschränkung auf Gespräche mit einem Kind wurde gewählt, weil bei Dialogen mit einer Kindergruppe nicht sichergestellt ist, dass die Kinder am kompletten Gespräch beteiligt sind und dementsprechend wirklich vom Dialog profitieren. So könnte ein Kind auf die erste Frage einer Fachkraft antworten, bei der sich anschließenden Frage

aber bereits mit etwas anderem gedanklich beschäftigt sein. Es ist jedoch auch zu überlegen, ob viele Einzelgespräche während einer Bilderbuchbetrachtung wirklich sinnvoll und erstrebenswert sind. Wenn die Bilderbuchbetrachtung mit einer Gruppe von Kindern durchgeführt wird, könnten viele längere Gespräche mit einzelnen Kindern auch dazu führen, dass manche Kinder weniger beteiligt sind und daraufhin gegebenenfalls auch das Interesse verlieren. Um diese Fragen beantworten zu können, müssten die Wirkungen von längeren Gesprächen genauer untersucht werden.

In Bezug auf das Wiederholen und Erweitern ist festzustellen, dass die Äußerungen der Kinder im Durchschnitt eher wiederholt als erweitert werden. Die syntaktische oder semantische Ergänzung von kindlichen Aussagen kommt nur ca. halb so oft vor wie die reine Wiederholung. Auch in der Studie von Girolametto und Weitzman (2002) wurden inhaltliche oder syntaktische Erweiterungen von den fröhpädagogischen Fachkräften eher selten eingesetzt. Vor allem für Kinder mit guten sprachlichen Fähigkeiten könnten jedoch gerade die zusätzlichen Informationen hilfreicher sein als reine Wiederholungen. Dies sollte daher noch stärker in den Fokus von Qualifizierungen gerückt werden. Korrektive Wiederholungen oder Erweiterungen wurden deutlich seltener beobachtet als die nicht korrektiven Formen. Da nur „fehlerhafte“ Äußerungen von Kindern korrigiert werden können, ist die geringere Auftretenshäufigkeit nachvollziehbar. Eine in diesem Zusammenhang interessante Fragestellung für weitere Untersuchungen könnte sein, welcher Anteil an „fehlerhaften“ Äußerungen von Fachkräften korrigierend aufgegriffen wird.

### *Effekte des Feedbackgesprächs*

Wie bereits bei der globalen Einschätzung der sprachlichen Anregung, wurden auch für die einzelnen sprachlichen Verhaltensweisen zu T1 keine signifikanten Unterschiede zwischen der Kontroll- und der Feedbackgruppe gefunden. Es ist daher von einem ähnlichen Ausgangsniveau auszugehen. Was die Veränderung zu T2 betrifft, sind jedoch wesentliche Unterschiede zu verzeichnen.

Ausgehend von den im Feedback thematisierten Inhalten wurde für die Fragen oder Erzählaufforderungen mit Antwortmöglichkeit ein positiver Effekt des Feedbacks erwartet. Sowohl für die Ja/Nein-Fragen als auch für die geschlossenen und die offenen Fragen sowie die Erzählaufforderungen mit Antwortmöglichkeit ergab die statistische Auswertung einen signifikanten Interaktionseffekt zwischen Messzeitpunkt und Bedingung. Bezüglich aller vier Variablen zeigten sich bei den Fachkräften der Kontrollbedingung keine signifikanten Veränderungen in den *t*-Tests für abhängige Stichproben. Die Fachkräfte, die am online-basierten Videofeedback teilgenommen haben, konnten sich hingegen zu T2 signifikant verbessern. Bei den offenen Fragen mit Antwortmöglichkeit entspricht der Unterschied in der Differenz zwischen T1 und T2 einem großen Effekt. Die Wirkung auf die An-

zahl an gestellten offenen Fragen ist vergleichbar mit den Befunden der Studien von Milburn et al. (2014) und Wasik et al. (2006), wobei die Qualifizierung der Fachkräfte dort jeweils aus mehreren Workshop-Tagen und Coaching-Sitzungen bestand und damit deutlich aufwendiger waren. Für das „Heidelberger Interaktionstraining“ konnten von Simon und Sachse (2011) keine Effekte auf den Einsatz von offenen Fragen nachgewiesen werden. Betrachtet man zusätzlich die in der Literatur angenommene besonders sprachförderliche Wirkung von offenen Fragen (de Rivera et al., 2005; Dickinson & Tabors, 2001), ist der hier beobachtete Effekt als äußerst positiv zu bewerten. In Bezug auf die Ja/Nein-Fragen, die geschlossenen Fragen und die Erzählaufforderungen mit Antwortmöglichkeit weisen die Feedbackeffekte eine mittlere bis große Stärke auf. Dies ist konform mit den Ergebnissen von Milburn et al. (2014), in deren Studie ebenfalls nach der Intervention mehr geschlossene Fragen bei der Bilderbuchbetrachtung gestellt wurden. Auf den ersten Blick mag es angesichts der eingeschränkten Antwort- und Beteiligungsmöglichkeiten dabei nicht positiv erscheinen, dass die Fachkräfte nach dem Feedback mehr Ja/Nein-Fragen und geschlossene Fragen verwenden. Sowohl bei den Kindern als auch bei den Fachkräften muss allerdings in der „Zone der nächsten Entwicklung“ gedacht werden. Was die Kinder betrifft, können insbesondere ruhigere Kinder oder Sprachanfänger durch geschlossene Fragen oder Ja/Nein-Fragen im Sinne des „scaffolding“ die Gelegenheit erhalten, sich am Gespräch zu beteiligen (Blewitt et al., 2009; Pentimonti & Justice, 2010). Mit Blick auf die „Zone der nächsten Entwicklung“ der Fachkräfte wurde in den Feedbackgesprächen häufig allgemein das Stellen von Fragen thematisiert und anvisiert, vor allem wenn zu T1 fast ausschließlich vorgelesen wurde. In der Folge wird auch die gestiegene Anzahl an Ja/Nein-Fragen oder geschlossenen Fragen in der Feedbackgruppe als positiver Effekt des Feedbacks gedeutet. Dies gilt vor allem vor dem Hintergrund, dass gleichzeitig auch mehr offene Fragen eingesetzt wurden. Da die Kombination aus weniger und stärker fordernden Anregungen besonders lernunterstützend sein kann (Blewitt et al., 2009), sollte das Zusammenspiel von geschlossenen und offenen Fragen in der Zukunft noch weiter erforscht werden.

Da bei Fragen oder Aufforderungen ohne Antwortmöglichkeit sowie vermeintlichen Fragen für die Kinder nicht von einem hohen sprachanregenden Charakter auszugehen ist, wurden diese in den Feedbackgesprächen nicht als positives Zielverhalten, sondern eher als zu vermeidendes Verhalten definiert. Erwartet wurde daher, dass entweder keine Veränderungen oder eine Abnahme in der Häufigkeit festgestellt werden. Bei den geschlossenen Fragen und den Ja/Nein-Fragen ohne Antwortmöglichkeit sowie den vermeintlichen Fragen liegt kein signifikanter Interaktionseffekt vor. Das Feedback hatte demnach keinen Einfluss auf dieses Verhalten. Hinsichtlich der offenen Fragen ohne Antwortmöglichkeit fand

entgegen der Erwartung eine signifikante Steigerung in der Feedbackgruppe statt, während in der Kontrollgruppe keine bedeutenden Veränderungen zu verzeichnen sind. Eine mögliche Erklärung dafür könnte sein, dass das bewusste Einbringen von offenen Fragen verhältnismäßig eine größere Herausforderung darstellt und es daher schwer ist, zusätzlich die Antwortmöglichkeit für die Kinder mit zu berücksichtigen. Darüber hinaus ist es bei offenen Fragen deutlich schwieriger, den Kindern ausreichend Zeit zu geben und eine Antwort abzuwarten als bei geschlossenen Fragen oder Ja/Nein-Fragen. Angesichts der Ergebnisse zur Antwortmöglichkeit liegt die Vermutung nahe, dass dieser Aspekt entweder nicht intensiv genug thematisiert wurde oder ein einziges Feedbackgespräch nicht ausreicht, um zu bewirken, dass die Fachkräfte die Kinder stärker durch Fragen einbeziehen und gleichzeitig darauf achten, den Kindern bei jeder Frage die Möglichkeit zum Antworten zu geben. Dementsprechend sollte dieses Thema beim Feedback stärker in den Fokus genommen werden. Wenn die Ressourcen es zulassen, wäre auch zu überlegen, diesen Aspekt in einem zweiten Feedbackgespräch zu reflektieren.

Sowohl für korrektive als auch für nicht korrektive Formen der Wiederholung und Erweiterung konnte eine signifikant positive Wirkung des Feedbacks konstatiert werden. Dabei handelt es sich im Falle der Wiederholungen und Erweiterungen ohne Korrektiv um große Effekte und bei den korrektiven Varianten um kleine bis mittelgroße Unterschiede zwischen Kontroll- und Feedbackgruppe. Während sich die Kontrollgruppe in keiner der Variablen bedeutend verändert hat, kann in der Feedbackgruppe bei den korrektiven Formen von kleineren bis mittleren und bei den nicht korrektiven Formen von mittleren bis großen Verbesserungen ausgegangen werden. Die unterschiedlich hohen Effektstärken sind insofern nicht verwunderlich, als die korrektive Wiederholung oder Erweiterung immer auf „fehlerhaften“ Äußerungen der Kinder basiert und daher von den Fachkräften nur bedingt beeinflussbar ist.

Über Fragen und Wiederholungen hinaus kann die Fachkraft auch über bezugnehmende Äußerungen an die Aussagen oder Fragen der Kinder anknüpfen. Da dies ebenfalls eine gute Möglichkeit ist, einen Dialog aufzubauen, wurde es in den Feedbackgesprächen unter dem Indikator „Gespräche“ diskutiert. Mit der ersten Hypothese konform stellte sich heraus, dass sich nur die Fachkräfte in der Feedbackbedingung signifikant verbessert haben. Milburn et al. (2014) haben in ihrer Studie die Aspekte Wiederholung, Erweiterung und Anknüpfen in der Variable „responsive comments“ zusammengefasst und signifikante Interventionseffekte beobachtet. In ihrem Fall konnte insgesamt ein großer Effekt erzielt werden. Ebenfalls positive Effekte stellten Simon und Sachse (2011) für korrektives Feedback fest. Die gefundenen Effekte replizieren daher die Ergebnisse ähnlicher Studien.

Inwiefern die positiven Veränderungen bezüglich der Fragen oder Kommentare auch zu längeren Dialogen mit einzelnen Kindern geführt haben, war eine weitere Fragestellung dieser Arbeit. Der Hypothese widersprechend, zeigte sich kein Effekt des Feedbacks auf die Anzahl an längeren Gesprächen. Ein denkbarer Grund könnte darin liegen, dass die Bilderbuchbetrachtungen in der vorliegenden Studie in der Regel mit einer Kindergruppe stattgefunden haben und die Fachkräfte, wie oben diskutiert, eher gruppenbezogen agieren und ihren Fokus weniger auf Gespräche mit einzelnen Kindern legen. In Bezug auf die Anzahl an längeren Gesprächen mit einem oder mehreren Kindern fanden Milburn et al. (2014) positive Effekte ihres Trainings.

Die Tatsache, dass das handlungsbegleitende Sprechen während der Feedbackgespräche kaum thematisiert wurde, findet sich in den erwartungsgemäß nicht signifikanten Feedbackeffekten für die Variablen „Selftalk“ und „Paralleltalk“ wieder.

Durch die Teilnahme am Feedbackgespräch konnten sich die Fachkräfte in dieser Studie signifikant in ihrem Frageverhalten sowie dem Einsatz von sprachmodellierenden Verhaltensweisen in Form von Wiederholungen und Erweiterungen der kindlichen Äußerungen verbessern. In der Feedbackforschung konkurrierten lange Zeit verschiedene Ansichten zur Funktion von Feedback. Vertreter der behavioristischen Feedbackforschung betonten die verstärkende Wirkung von Feedback, die kognitivistische Perspektive betrachtete Feedback in erster Linie als Information zur Korrektur von Fehlern. Narciss (2006) kommt zu dem Schluss, dass eine positive Wirkung von Feedback sowohl auf der Verstärkung von richtigen Antworten als auch auf der Korrektur falscher Lösungen basiert. Die Ergebnisse dieser Studie lassen sich mit dieser Annahme in Einklang bringen. Das Wiederholen und Erweitern wurde in den Feedbackgesprächen zum größten Teil als erster Indikator thematisiert, bei dem nur gelungene Verhaltensbeispiele zurückgemeldet wurden. Da für alle Formen der Wiederholung und Erweiterung signifikante Feedbackeffekte festgestellt wurden, muss eine positive Verstärkung dieser Verhaltensweisen stattgefunden haben. Das Stellen von offenen Fragen wurde hingegen fast ausschließlich als zweiter oder dritter Indikator reflektiert. Dabei lag der Fokus auf dem Optimierungspotenzial, weshalb die große Zunahme nach dem Feedbackgespräch auch für eine korrigierende Wirkungsweise spricht. Wie bereits erwähnt, wurde die Tatsache, dass Stärken und Schwächen reflektiert wurden, zugleich von einigen Fachkräften sehr positiv bewertet.

Zur Überprüfung der Effekte auf die einzelnen sprachlichen Verhaltensweisen der Fachkräfte wurde deren absolute Häufigkeit durch Auszählen für jede Videoaufnahme ermittelt. Um den unterschiedlichen Möglichkeiten in der Gestaltung von Bilderbuchbetrachtungen gerecht zu werden, wurde immer die gesamte Bilderbuchbetrachtung berücksichtigt. Es hat sich herausgestellt, dass sich die Länge der Bilderbuchbetrachtung zu T1 nicht zwischen

Kontroll- und Feedbackgruppe unterscheidet, wohingegen zu T2 in der Feedbackbedingung die Bilderbuchbetrachtungen signifikant länger gedauert haben. An dieser Stelle könnte man schlussfolgern, dass die zu T2 in der Feedbackgruppe beobachtete höhere Anzahl an Fragen, Erzählaufforderungen sowie Wiederholungen und Erweiterungen auf eine längere Dauer der Bilderbuchbetrachtungen und nicht auf das Feedback zurückzuführen ist. Da die Fachkräfte zu T2 das gleiche oder ein sehr ähnliches Buch verwendet haben, erscheint diese Kausalrichtung nicht logisch. Vielmehr wird davon ausgegangen, dass die Fachkräfte in der Feedbackgruppe zum zweiten Messzeitpunkt die Kinder stärker einbezogen und zum Sprechen angeregt haben, weshalb die Bilderbuchbetrachtungen länger gedauert haben. Die signifikanten Unterschiede in der Länge der Bilderbuchbetrachtungen zu T2 würden demnach die beobachteten Feedbackeffekte eher stützen. Dafür spricht zudem, dass sich die positiven Feedbackeffekte auch in der Dimension „Language Modeling“ der CLASS gezeigt haben, welche die einzelnen Verhaltensweisen in einer allgemeinen Einschätzung der sprachlichen Anregung bündelt und zeitunabhängig ausgewertet wird.

Betrachtet man den Durchschnitt der teilnehmenden Fachkräfte, kann das Feedback aufgrund der signifikant positiven Wirkung auf die meisten Variablen der sprachlichen Anregung als eine erfolgreiche Intervention betrachtet werden. Das online-basierte Videofeedback entspricht somit der Forderung, dass Trainings die performative Struktur von Kompetenzen und somit den Transfer unterstützen sollten (Fröhlich-Gildhoff, 2014).

### Individuelle Veränderungen

Um Aussagen über die individuelle Wirksamkeit treffen zu können, wurde mithilfe des „Reliable Change Index“ nach Jacobson und Truax (1991) berechnet, bei wie vielen Fachkräften eine signifikante Verbesserung stattgefunden hat.

Im handlungsbegleitenden Sprechen, den längeren Dialogen, den vermeintlichen Fragen sowie den Fragen ohne Antwortmöglichkeit haben relativ wenige positive Veränderungen stattgefunden, was den Ergebnissen der Varianzanalysen entspricht. Negative Veränderungen kamen insgesamt eher selten vor. Angesichts der Tatsache, dass einige Studien zur Wirksamkeit von Trainings im frühpädagogischen Kontext (Egert, 2015) und ca. ein Drittel der in Meta-Analysen einbezogenen Feedback-Interventionen (Kluger & DeNisi, 1996) insgesamt von negativen Effekten berichten, wird dieses Ergebnis sehr positiv interpretiert.

Bei einem Konfidenzintervall von 95 % ist bei weniger als 20 % der Fachkräfte von einer reliablen positiven Veränderung im „Language Modeling“ auszugehen. Grund dafür ist, dass die Standardabweichung in der Variable „Language Modeling“ in der vorliegenden Studie relativ hoch ist und somit auf dem 5 %-Niveau alle Veränderungen kleiner als 1.17 nicht als reliabel angesehen werden. Verbesserungen um einen Punkt auf der siebenstufigen Skala fallen daher nach dieser Berechnung nicht unter positive Veränderungen. Ein



erweitertes Konfidenzintervall von 10 % schließt diese Fälle mit ein. Nach dieser Definition können für fast zwei Drittel der Fachkräfte bedeutende Verbesserungen bescheinigt werden. Da auch in anderen umfangreicheren Interventionen in der Regel durchschnittlich nur Verbesserungen von einem Punkt oder weniger erreicht werden (z.B. Zan & Donegan-Ritter, 2014), kann ein derartiges Ergebnis nur als sehr positiv gewertet werden. Ein ähnliches Phänomen liegt bei den offenen Fragen mit Antwortmöglichkeit vor. Mit einer Sicherheitswahrscheinlichkeit von 5 % kann die positive Veränderung bei etwas mehr als einem Drittel der Fachkräfte als reliabel eingestuft werden. Bei einem Konfidenzintervall von 90 % werden für 50 % der Fachkräfte positive Veränderungen konstatiert.

Bezüglich der Ja/Nein-Fragen mit Antwortmöglichkeit wurden verhältnismäßig viele positive und negative Veränderungen festgestellt. Dies könnte damit zusammenhängen, dass das Feedback immer bei der Zone der nächsten Entwicklung angesetzt hat- je nach individuellen Voraussetzungen der Fachkraft kann eine Abnahme oder Zunahme an Ja/Nein-Fragen sinnvoll sein. Eine qualitative Einzelfallanalyse könnte hierzu weiter Aufschluss geben, war jedoch nicht Gegenstand dieser Arbeit. Fast die Hälfte der Fachkräfte in der Feedbackbedingung hat zu T2 bedeutend mehr geschlossene Fragen gestellt. Bei den Wiederholungen ohne Korrektiv zeigte ebenfalls ca. die Hälfte der Fachkräfte bedeutende positive Veränderungen. Gemäß den varianzanalytischen Gesamtergebnissen fällt die Anzahl an Verbesserungen bei den korrektiven Formen niedriger aus.

Auf den ersten Blick mögen die Befunde nicht für die Effektivität des Feedbacks sprechen, sind jedoch vergleichbar mit bisherigen Befunden. Jacobson und Truax (1991) berichten beispielsweise für klinische Studien Ergebnisse in einer ähnlichen Größenordnung. In der pädagogischen Interventionsforschung wurde die individuelle Wirksamkeit von Trainings bislang kaum untersucht. Die Studie von Girolametto et al. (2007) ist eine der wenigen Untersuchungen, welche auf individuelle Effekte eingeht. Sie konnten nach der Teilnahme an einem, auf zwei Tage gekürzten, Training aus dem „Hanen-Program“ bei ca. der Hälfte der teilnehmenden Fachkräfte eine Veränderung in den erfassten Variablen beobachten. Dies entspricht damit in etwa den hier gefundenen Verbesserungen. Allerdings basieren die Ergebnisse von Girolametto et al. (2007) auf der Analyse der Veränderungen von nur acht Fachkräften. Zudem wurde für die Auswertung ein von den Autoren festgelegtes Kriterium ohne Signifikanzprüfung verwendet. Angesichts der Ergebnisse anderer Studien und der für das Feedback erforderlichen geringen zeitlichen Ressourcen erscheinen die ermittelten Verbesserungen durchaus positiv. Da das Ziel einer Intervention immer sein sollte, dass so viele Teilnehmer wie möglich davon profitieren, sollte die individuelle Wirksamkeit in zukünftigen Studien unbedingt stärker berücksichtigt werden.

### Einfluss von anderen Variablen

Um sicherzustellen, dass die Effekte des Feedbackgesprächs nicht auf anderen Aspekten basieren, wurde geprüft, ob sich die beiden Gruppen in relevanten Merkmalen unterscheiden (Merkmale der Fachkraft, situative Unterschiede zwischen T1 und T2, Leistung und Motivation zu T1). Die Ergebnisse der *t*-Tests und Chi-Quadrat-Tests ergaben signifikante Unterschiede in Bezug auf den Migrationshintergrund, die Berufserfahrung als Fachkraft sowie die Differenz in der Anzahl der Kinder zwischen T1 und T2. Hinsichtlich dieser Variablen konnte kein Einfluss auf die Effekte des Feedbacks gefunden werden. Die Effekte des Feedbacks sind daher insgesamt nicht auf bestehende Unterschiede zwischen der Kontrollgruppe und der Feedbackgruppe zurückzuführen.

In hierarchischen Regressionsanalysen wurde zusätzlich der Einfluss von verschiedenen Variablen (Selbstwirksamkeit und Einstellung gegenüber Feedback, Zielorientierung, Berufsbezogene Motivation, Reflexionsniveau der Fachkraft, Bewertung des Feedbacks durch die Fachkräfte und technische Probleme während des Feedbacks) auf die Veränderungen innerhalb der Feedbackgruppe analysiert. Dabei wurden immer im ersten Hierarchieschritt die jeweiligen Störvariablen als Prädiktoren aufgenommen. Insgesamt ist anzumerken, dass die Leistung zu T1 der stärkste Prädiktor war. Bei Fachkräften mit einer höheren Leistung zu T1 fallen häufig die Verbesserungen geringer aus. Dies ist gut nachvollziehbar, da sich Fachkräfte mit einer sehr guten Leistung zu T1 nicht mehr so viel verbessern können (Deckeneffekt). Die Regressionsanalysen liefern Hinweise darauf, dass außerdem Fachkräfte mit einem höheren Abschluss und einer höheren Funktion in manchen Aspekten eher vom Feedback profitiert haben. Dabei muss beachtet werden, dass die Regressionsgewichte dieser Störvariablen unter Berücksichtigung der Anfangsleistung häufig keine statistische Bedeutsamkeit erreichten. Im zweiten Regressionsmodell wurden die jeweiligen potenziellen Einflussvariablen zusätzlich zu den Störvariablen als Prädiktoren berücksichtigt. Da nur sehr vereinzelt die Differenzwerte der sprachlichen Anregung durch die jeweiligen Variablen vorhergesagt wurden, kann insgesamt nicht von einem bedeutenden Zusammenhang zwischen den einzelnen Variablen und der Wirkung des Feedbacks ausgegangen werden. Die Tatsache, dass technische Probleme während des Feedbacks scheinbar keinen Einfluss auf die Effekte des Feedbacks hatten, ist positiv zu bewerten. Die Ergebnisse widersprechen den Befunden von Studien, die auf einen bedeutenden Einfluss der Selbstwirksamkeit (Heslin & Latham, 2004) und der Lernzielorientierung (Payne et al., 2007) auf die Effekte von Feedback hinweisen. Die Motivation der Fachkräfte war bislang zwar eher selten Gegenstand von Feedbackuntersuchungen, aber Erkenntnisse aus der allgemeinen Lehr-/Lernforschung liefern durchaus Belege für eine Verbindung zwischen Motivation und Lernleistung (Schiefele & Schreyer, 1994). Um belastbare Aussagen treffen zu können, sollten

diese Zusammenhänge weiter untersucht werden. So könnte eine noch größere Stichprobe die Gefahr minimieren, dass bedeutsame Regressionsgewichte nicht signifikant werden (Bühner & Ziegler, 2017). Auch eine Modifizierung in der Erfassung der Variablen wäre denkbar. Vor allem bei leistungsbezogenen Aussagen zum Vergleich mit KollegInnen im LMI-K und beim AGQ-R haben sich die teilnehmenden Fachkräfte häufig mit der Beantwortung schwer getan. In Bezug auf das Reflexionsniveau der Fachkräfte könnte eine differenziertere Erfassung hilfreich sein.

### 8.3 Grenzen der vorliegenden Studie

Die vorliegende Studie konnte positive Effekte eines online-basierten Videofeedbacks auf das sprachliche Interaktionsverhalten von Fachkräften aufzeigen. Da keine Follow-up-Messung stattgefunden hat, können keine Aussagen zur Nachhaltigkeit der Effekte gemacht werden. Dies sollte in weiteren Studien nachgeholt werden. Der Einbezug von einer größeren Anzahl an videografierten Situationen könnten zudem ein noch vollständigeres Bild zum Verhalten der Fachkräfte geben.

Die Ergebnisse der Studie sprechen sowohl für eine verstärkende als auch eine korrigierende bzw. optimierende Wirkung des Feedbacks. Die in der Literatur angenommene motivierende Funktion von Feedback (siehe Kapitel 3.2.1 und Kapitel 3.2.2) wurde nicht überprüft. Untersuchungen zu den Effekten eines solchen Feedbacks auf die Motivation der Fachkräfte stehen demnach noch aus.

Darüber hinaus liefert diese Studie keine Ergebnisse zu den mittelbaren Effekten des Feedbacks für die frühpädagogischen Fachkräfte auf die Entwicklung der Kinder. In einigen Studien konnte der Zusammenhang zwischen einer hohen Qualität der mit der CLASS eingeschätzten Interaktionsqualität und einer positiven Entwicklung der Kinder nachgewiesen werden (z.B. Curby et al., 2009; Mashburn et al., 2008). Dementsprechend sollten sich Verbesserungen in der sprachlichen Anregung der Fachkräfte auch in der Sprache der Kinder äußern. Sowohl Egert (2015) als auch Markussen-Brown et al. (2017) fanden im Durchschnitt kleine Effekte der Trainings auf kindliche Entwicklungsmaße. Simon und Sachse (2013) untersuchten die Auswirkungen des „Heidelberger Interaktionstrainings“ auf die Sprechfreude und die verbalen Äußerungen der Kinder. Sie stellten fest, dass über eine Verbesserung der sprachlichen Anregung der Fachkräfte eine erhöhte Sprechfreude bei den Kindern und eine größere Komplexität ihrer Äußerungen erreicht werden konnte. Da letztlich jede Intervention zur Verbesserung der Interaktionsqualität darauf abzielt, die Entwicklung der Kinder zu unterstützen, sollte dies unbedingt in weiteren Forschungsvorhaben Berücksichtigung finden. Das Zusammenspiel mit den verschiedenen Anregungsformen stellt hierbei eine zusätzliche interessante Fragestellung dar. Beispielsweise könnte die Äu-

Berungslänge der Kinder sowie die Redeanteile in Abhängigkeit vom Einsatz der verschiedenen Fragetypen näher untersucht werden.

Wie bereits erwähnt, handelt es sich um eine selektive Stichprobe. Aufgrund der freiwilligen Teilnahme, kann angenommen werden, dass die teilnehmenden Fachkräfte prinzipiell gegenüber Feedback und der Arbeit mit Videoaufnahmen offen eingestellt waren. Hinsichtlich der Einstellung gegenüber Feedback kann dies durch die hohen Werte in den Skalen der „Feedback Orientation Scale“ (Linderbaum & Levy, 2010) bestätigt werden. Dies könnte die Effekte des Feedbacks beeinflusst haben. Die Studie hat zudem ausschließlich mit Fachkräften aus München und der Umgebung stattgefunden. Da jedoch keine regionalen Unterschiede hinsichtlich der Effekte von Trainings bekannt sind, ist diesbezüglich von einer Generalisierbarkeit der Befunde auszugehen.

Die Frage der Übertragbarkeit stellt sich auch in Bezug auf die inhaltliche Dimension und die Situation. In der dargestellten Studie wurden die Effekte auf die sprachliche Interaktion von Fachkräften während einer Bilderbuchbetrachtung geprüft. In künftigen Studien sollten die Feedbackeffekte auch in anderen Alltagssituationen untersucht werden. Da das Feedback-Konzept auf andere Dimensionen des Interaktionsverhaltens übertragen werden kann, wären diesbezügliche Evaluationsvorhaben ebenfalls von Interesse.

Der Fokus dieser Arbeit lag auf der allgemeinen Wirksamkeit des Feedbacks. Bedingungsfaktoren für die Effekte des Feedbacks wurden hauptsächlich in Form von verschiedenen Merkmalen der Fachkräfte einbezogen. Der Einfluss des Feedbackgebers wurde in dieser Studie statistisch kontrolliert. Bei Egert (2015) sind die Effekte von Fortbildungen in den Studien höher, in denen das Training vom wissenschaftlichen Personal selbst durchgeführt wurde. Einzelne Feedbackuntersuchungen liefern außerdem Hinweise darauf, dass die wahrgenommene Kompetenz des Feedbackgebers eine Bedeutung für die Auswirkungen von Feedback hat (z.B. Raemdonck & Strijbos, 2013). Eine genauere Beschäftigung mit der Rolle des Feedbackgebers erscheint daher lohnenswert. Da die Feedbackgespräche in dieser Studie ausschließlich von wissenschaftlichen Mitarbeiterinnen des Instituts für Frühpädagogik durchgeführt wurden, sollte mit Blick auf die Praxistauglichkeit getestet werden, ob sich die beobachteten Feedbackeffekte auch bei anderen Fortbildnern replizieren lassen.

## 9 Fazit und Ausblick

Im Rahmen dieser Arbeit wurde ein Konzept für online-basiertes Videofeedback, mit dem frühpädagogischen Fachkräften Rückmeldung zu ihrem sprachlichen Interaktionsverhalten gegeben wird, erprobt und wissenschaftlich überprüft.

Die Durchführung der Feedbackgespräche mit dem Leitfaden und der dazugehörigen Arbeitsvorlage hat sich als sehr praktikabel erwiesen. Sowohl die inhaltliche als auch die technische Umsetzung haben insgesamt gut funktioniert. Auch die teilnehmenden Fachkräfte waren sehr zufrieden mit dem Feedback. Neben der gesamten Struktur und dem Aufbau des Feedbacks wurde von den Fachkräften vor allem die Arbeit mit den Videosequenzen sehr positiv bewertet. Einige Fachkräfte haben es dabei als hilfreich empfunden, dass sowohl Stärken als auch Entwicklungspotenziale thematisiert wurden und betonten die wertschätzende und freundliche Art der Kommunikation als gelungene Aspekte des Feedbacks. Zeitersparnis und zeitliche Flexibilität wurden als Vorteile der internet-basierten Umsetzung gesehen.

Da es jedoch in vielen Feedbackgesprächen zu vereinzelten technischen Schwierigkeiten kam, ist diesbezüglich noch Optimierungspotenzial vorhanden. Um ein solches online-basiertes Feedback in Zukunft sinnvoll als Instrument der Qualitätsentwicklung einsetzen zu können, müssen pädagogische Fachkräfte Zugang zu einem gut funktionierenden Computer mit den nötigen Nutzungsrechten sowie einer stabilen Internetverbindung haben. Die teilnehmenden Fachkräfte haben dies in ihrer Einschätzung des Feedbackgesprächs wie folgt zusammengefasst: Diese Form des Feedbacks stellt eine gute Möglichkeit der Fortbildung dar, sofern die Technik funktioniert.

Die Erfahrungen in dieser Studie sowie in anderen Untersuchungen (z.B. Werner et al., 2018) haben gezeigt, dass sich Feedbackgespräche von einer Stunde gut in den Alltag der Fachkräfte integrieren lassen. Auch die Videoaufnahmen konnten in der Regel ohne großen zusätzlichen Aufwand für die Fachkräfte angefertigt werden. Trotzdem haben einige Fachkräfte die Teilnahme an der Studie wegen mangelnder Zeit abgesagt. Eine wichtige Voraussetzung für eine mögliche Implementierung des Feedbacks ist daher, dass ausreichend Zeit für Qualitätsentwicklungsmaßnahmen zur Verfügung gestellt wird.

Ein weiterer Grund, nicht an der Untersuchung teilzunehmen, war für einige Fachkräfte die Arbeit mit Videoaufnahmen. Gleichzeitig hat den teilnehmenden Fachkräften gerade die Reflexion mithilfe der Videosequenzen am besten gefallen. Eine umfangreiche Aufklärungsarbeit bei Fachkräften und Eltern zu den Potenzialen von Videoaufnahmen und zum Thema Datenschutz könnten diesem Problem entgegenwirken. Beispielsweise könnten

Fachkräfte ihre positiven Erfahrungen mit solchen Formaten mit anderen Fachkräften teilen und sie so dazu ermutigen.

Auf Basis der Erkenntnisse dieser Studie lassen sich demnach die drei folgenden wichtigen Gelingensbedingungen für eine Implementierung des online-basierten Videofeedbacks festhalten: Eine gute technische Ausstattung in den Kitas, eine offene Einstellung aller Beteiligten gegenüber online-basiertem Videofeedback sowie ausreichend Zeit für die Teilnahme an Fortbildungen.

Auch auf Seiten des Feedbackgebers gibt es wesentliche Voraussetzungen für die erfolgreiche Durchführung von Feedbackgesprächen. Für jedes Feedbackgespräch ist zunächst eine Einschätzung des Ist-Stands notwendig. Die Erfahrungen in dieser Studie haben einmal mehr gezeigt, dass eine reliable und objektive Ermittlung des Ist-Stands nur mithilfe einer guten Schulung in Bezug auf die relevanten Verhaltensaspekte und deren Erfassung möglich ist. Im Einklang damit weisen wissenschaftliche Ergebnisse darauf hin, dass Feedback effektiver ist, wenn es von geschulten Beobachtern gegeben wird (z.B. Feys et al., 2011). Basierend auf der Einschätzung des Verhaltens wurden die Feedbackgespräche anhand des entwickelten Konzepts durchgeführt. Dabei wurde mit einem Leitfaden sowie einer darauf aufbauenden Arbeitsvorlage gearbeitet. Feedbackgeber sollten gut mit dem Konzept und den dazugehörigen Materialien vertraut sein. Die Notwendigkeit dessen wird auch empirisch gestützt, da Trainings höhere Effekte haben, wenn sie mit einem festen Curriculum arbeiten (Fukkink & Lont, 2007). Hinsichtlich der Übermittlung des Feedbacks scheint für die Fachkräfte zudem die Art der Kommunikation eine wichtige Rolle zu spielen. Die Qualifizierung von Feedbackgebern sollte demzufolge eine Schulung zur reliablen Beurteilung des Verhaltens der Fachkräfte sowie eine intensive Auseinandersetzung mit dem Curriculum und wichtigen Aspekten der Gesprächsführung beinhalten.

Zur Überprüfung der Effekte des Feedbacks auf das sprachliche Interaktionsverhalten der Fachkräfte wurde zu zwei Messzeitpunkten jeweils eine Videoaufnahme während einer Bilderbuchbetrachtung bei jeder Fachkraft gemacht. Die Auswertung dieser Videoaufnahmen mithilfe der Dimension „Language Modeling“ aus dem „Classroom Assessment Scoring System“ (CLASS-Pre-K; Pianta, La Paro et al., 2008) sowie einem zusätzlichen Kodier-System gibt einen wichtigen Einblick in die Qualität der sprachlichen Anregung bei Bilderbuchbetrachtungen. Im Durchschnitt erreichten die Fachkräfte einen Wert im unteren mittleren Niveau im „Language Modeling“. Dieser Befund spiegelt die Ergebnisse anderer Studien zur Qualität sprachlicher Anregung im Kindergarten (z.B. Suchodoletz et al., 2014; Wirts et al., 2017) wieder. Bezüglich einzelner Verhaltensweisen ergab die Auswertung unter anderem, dass die Fachkräfte deutlich mehr geschlossene Fragen und Ja/Nein-Fragen stellen als offene Fragen und reine Wiederholungen kindlicher Äußerungen viel häufiger vorkom-

men als zusätzliche Erweiterungen. Auch dies entspricht den Befunden vorangegangener Studien (Girolametto & Weitzman, 2002; Milburn et al., 2014; Simon & Sachse, 2011). Für die sprachliche Anregung von Kindern im Alter zwischen drei und sechs Jahren im Rahmen von Bilderbuchbetrachtungen besteht daher nach wie vor ein großer Entwicklungsbedarf. Bezüglich der meisten der erfassten Variablen konnten durch das Feedback signifikante Verbesserungen erzielt werden. Die Auswertung ergab, dass in der Kontrollgruppe keine Veränderungen von T1 zu T2 stattgefunden haben. Die Feedbackgruppe zeigte hingegen zu T2 durchschnittlich höhere Werte im „Language Modeling“, eine größere Anzahl an Fragen und Erzählaufforderungen mit Antwortmöglichkeit sowie mehr Wiederholungen und Erweiterungen als zu T1. Die größte Mittelwertsteigerung konnte für die offenen Fragen mit Antwortmöglichkeit festgestellt werden. Ca. zwei Drittel der Fachkräfte in der Feedbackgruppe haben sich im „Language Modeling“ um einen Punkt auf der sieben-stufigen Skala verbessert. Mit Blick auf die einzelnen sprachlichen Verhaltensweisen wurden die meisten individuellen positiven Veränderungen in den Fragen mit Antwortmöglichkeit sowie den Wiederholungen ohne Korrektiv beobachtet. Angesichts der im Feedback thematisierten Inhalte sprechen die Ergebnisse für eine verstärkende und eine korrigierende Wirkung des Feedbacks. Für die künftige Anwendung von Feedback im pädagogischen Kontext empfiehlt sich daher die Berücksichtigung beider Facetten.

Die Effekte des Feedbacks wurden in einem Prä-Post-Kontrollgruppendesign mit einer ausreichend großen Stichprobe auf Verhaltensebene unter Berücksichtigung individueller Unterschiede getestet. Damit sind wichtige methodische Anforderungen an Evaluationsstudien erfüllt, die in der bisherigen Forschung häufig vernachlässigt wurden (Kluger & DeNisi, 1996). Einschränkungen der Studie und damit einhergehender Forschungsbedarf liegen vor allem in der Nachhaltigkeit der Effekte, der Erfassung der Auswirkungen auf kindlicher Ebene und der Übertragbarkeit auf andere Verhaltensaspekte und Situationen.

Zusammenfassend kann aus den Ergebnissen abgeleitet werden, dass sich die Inanspruchnahme des vorgestellten online-basierten Videofeedbacks für frühpädagogische Fachkräfte lohnt. Die auf dem Feedback basierenden Veränderungen auf Fachkraft-Ebene entsprechen zum Teil großen Effekten und gewinnen umso mehr an Bedeutung, als es sich um eine einmalige Intervention von ca. einer Stunde gehandelt hat. Da für deutlich umfangreichere Programme im Durchschnitt vergleichbare Effekte gefunden wurden (Egert, 2015), kann beim hier durchgeführten Feedback von einer guten Kosten-Nutzen-Balance ausgegangen werden. Das Feedback ist daher nicht nur im Rahmen von individuellen Fortbildungen, sondern auch im Kontext von wissenschaftlichen Projekten oder breit angelegten Qualifizierungsoffensiven gut einsetzbar.

Fortbildungsformate mit individuellen Elementen, wie Feedback oder Coaching, haben sich insgesamt in verschiedenen Meta-Analysen als am effektivsten erwiesen (Markussen-Brown et al., 2017; Werner et al., 2016). Auf der Grundlage der Befunde der genannten Untersuchungen und den Ergebnissen dieser Studie wird die Empfehlung ausgesprochen, in Zukunft eher in die individuelle Unterstützung der Fachkräfte zu investieren als in traditionelle Formate mit Seminar-Charakter.

Mit den Erkenntnissen aus der Erprobung des online-basierten Videofeedbacks und der empirischen Überprüfung der Effekte auf Performanz-Ebene leistet diese Arbeit einen wichtigen Beitrag für die Qualitätsentwicklung im frühpädagogischen Kontext.



## 10 Literaturverzeichnis

- Aalsvoort, G. M. (2003). Kognitive Kompetenzverbesserung bei Vorschulkindern mit Schulerfolgsrisiko. *Psychologie in Erziehung und Unterricht*, 50, 199-209.
- Ahnert, L. (2004). Bindungsbeziehungen außerhalb der Familie. Tagesbetreuung und Erzieherinnen-Kind-Bindung. In L. Ahnert (Hrsg.), *Frühe Bindung- Entstehung und Entwicklung* (S. 256-277). München: Reinhardt.
- Ahnert, L., Pinquart, M. & Lamb, M. (2006). Security of Children's Relationships With Nonparental Care Providers: A Meta-Analysis. *Child Development*, 74(3), 664-679.
- Ahnert, L. (2010). *Wieviel Mutter braucht ein Kind? Bindung - Bildung - Betreuung: öffentlich und privat*. Berlin u.a.: Springer Spektrum.
- Ainsworth, M. D. S., Bell, S. M. & Stayton, D. J. (1974). Infant-mother attachment and social development: Socialization as a product of reciprocal responsiveness to signals. In M. J. M. Richards (Hrsg.), *The integration of a child into a social world* (S. 99-135). London: Cambridge University Press.
- Ainsworth, M. D. S., Blehar, M. C., Waters, E. & Wall, S. (1978). *Patterns of attachment. A psychological study of the strange situation*. Hillsdale, NJ: Erlbaum.
- Altermann, A. & Holmgaard, M. (2015). *Der Akademisierungsprozess im Arbeitsfeld Kita aus Sicht der Träger. Eine Befragung in Baden-Württemberg, Nordrhein-Westfalen und Thüringen* (Weiterbildungsinitiative Frühpädagogische Fachkräfte, Hrsg.) (WiFF Studien Nr. 26). München.
- Ammons, R. B. (1956). Effects of Knowledge of Performance. A Survey and Tentative Theoretical Formulation. *The Journal of General Psychology*, 54(2), 279-299.
- Anders, Y., Rossbach, H.-G., Weinert, S., Ebert, S., Kuger, S., Lehrl, S. et al. (2012). Home and preschool learning environments and their relations to the development of early numeracy skills. *Early Childhood Research Quarterly*, 27(2), 231-244.
- Anderson, R. C., Kulhavy, R. W. & Andre, T. (1971). Feedback procedures in programmed instruction. *Journal of Educational Psychology*, 62(2), 148-156.
- Antons, K. (2000). *Praxis der Gruppendynamik. Übungen und Techniken* (8., durchges. und erg. Aufl.). Göttingen u.a.: Hogrefe Verl. für Psychologie.
- Arnett, J. (1989). *Caregiver Interaction Scale*. FPG Child Development Institute, UNC Chapel Hill.
- Atwater, L. E. & Brett, J. F. (2005). Antecedents and consequences of reactions to developmental 360° feedback. *Journal of Vocational Behavior*, 66(3), 532-548.

- Bailey, C. S., Denham, S. A., Curby, T. W. & Bassett, H. H. (2016). Emotional and organizational supports for preschoolers' emotion regulation. Relations with school adjustment. *Emotion, 16*(2), 263-279.
- Baldwin, T. T. & Ford, J. K. (1988). Transfer of training. A review and directions for future research. *Personnel Psychology, 41* (1), 63-105.
- Balzer, W., Doherty, M. & O'Connor, R. (1989). Effects of cognitive feedback on performance. *Psychological Bulletin, 106* (3), 410-433.
- Bandura, A. (1979). *Sozial-kognitive Lerntheorie* (1. Aufl.). Stuttgart: Klett-Cotta.
- Bandura, A. (1977). Self-efficacy. Toward a unifying theory of behavioral change. *Psychological Review, 84* (2), 191-215.
- Bandura, A. (1993). Perceived Self-Efficacy in Cognitive Development and Functioning. *Educational Psychologist, 28* (2), 117.
- Bangert-Drowns, R., Kulik, C.-L. C., Kulik, J. & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of educational research, 61* (2), 213-238.
- Bardwell, R. (1981). Feedback: How Does It Function? *The Journal of Experimental Education, 50* (1), 4-9.
- Baumeister, K. & Grieser, A. (2011). *Berufsbegleitende Fort- und Weiterbildung frühpädagogischer Fachkräfte - Analyse der Programmangebote. Eine Studie der Weiterbildungsinitiative Frühpädagogische Fachkräfte (WiFF)*. München: Deutsches Jugendinstitut e.V. (DJI).
- Bell, S. T. & Arthur, W. (2008). Feedback acceptance in developmental assessment centers. The role of feedback message, participant personality, and affective response to the feedback session. *Journal of Organizational Behavior, 29* (5), 681-703.
- Beller, S. & Beller, E. K. (2009). *Abschlussbericht des Projekts Systematische sprachliche Anregung im Kindergartenalltag zur Erhöhung der Bildungschancen 4- und 5-jähriger Kinder aus sozial schwachen und Migrantenfamilien - ein Modell der pädagogischen Intervention*. Berlin: Ein gemeinsames Projekt der Internationalen Akademie für innovative Pädagogik und Ökonomie gGmbH und der Freien Universität Berlin.
- Bertelsmann Stiftung. (2015). *Trends der FBBE in Deutschland - zentrale Ergebnisse des Länderreports 2015*. Verfügbar unter <https://www.laendermonitor.de/downloads-presse/index.nc.html>
- Birney, R. C., Burdick, H. & Teevan, R. C. (1969). *Fear of failure*. New York: Van Nostrand.
- Bishop, C. D., Snyder, P. A. & Crow, R. E. (2015). Impact of Video Self-Monitoring With Graduated Training on Implementation of Embedded Instructional Learning Trials. *Topics in Early Childhood Special Education, 35* (3), 170-182.

- Blewitt, P., Rump, K. M., Shealy, S. E. & Cook, S. A. (2009). Shared book reading. When and how questions affect young children's word learning. *Journal of Educational Psychology*, 101 (2), 294-304.
- BMFSFJ. (2005). *Zwölfter Kinder und Jugendbericht. Bericht über die Lebenssituation junger Menschen und die Leistungen der Kinder- und Jugendhilfe in Deutschland*. Bundesministerium für Familie, Senioren, Frauen und Jugend. Zugriff am 08.08.2017. Verfügbar unter <https://www.bmfsfj.de/blob/112224/7376e6055bbcaf822ec30fc6ff72b287/12-kinder-und-jugendbericht-data.pdf>
- Bowlby, J. (2002). Bindung: Historische Wurzeln, theoretische Konzepte und klinische Relevanz. In G. Spangler & P. Zimmermann (Hrsg.), *Die Bindungstheorie. Grundlagen, Forschung und Anwendung* (4. Aufl., S. 17-26). Stuttgart: Klett-Cotta.
- Bowlby, J. (2010). *Bindung als sichere Basis. Grundlagen und Anwendung der Bindungstheorie* (2. Aufl.). München u.a.: Reinhardt.
- Briedigkeit, E. (2011). Institutionelle Überformung sprachlicher Handlungsmuster – Realisation von Fragetypen im Erzieherin-Kind(er)-Diskurs. *Empirische Pädagogik*, 25 (4), 499-517.
- Broekhuizen, M. L., Mokrova, I. L., Burchinal, M. R. & Garrett-Peters, P. T. (2016). Classroom Quality at Pre-kindergarten and Kindergarten and Children's Social Skills and Behavior Problems. *Early Childhood Research Quarterly*, 36, 212-222.
- Bruner, J. S. (2008). *Wie das Kind sprechen lernt*. Bern: Hans Huber.
- Brutus, S., Derayeh, M., Fletcher, C., Bailey, C., Velazquez, P., Shi, K. et al. (2006). Internationalization of multi-source feedback systems. A six-country exploratory analysis of 360-degree feedback. *The International Journal of Human Resource Management*, 17 (11), 1888-1906.
- Bühner, M. & Ziegler, M. (2017). *Statistik für Psychologen und Sozialwissenschaftler* (Pearson Studium, 2., aktualisierte und erweiterte Auflage). Hallbergmoos: Pearson.
- Burchinal, M. R., Peisner-Feinberg, E., Pianta, R. & Howes, C. (2002). Development of Academic Skills from Preschool Through Second Grade. Family and Classroom Predictors of Developmental Trajectories. *Journal of School Psychology*, 40 (5), 415-436.
- Buschmann, A., Simon, S., Jooss, B. & Sachse, S. (2010). Ein sprachbasiertes Interaktions-training für Erzieherinnen ("Heidelberger Trainingsprogramm") zur alltagsintegrierten Sprachförderung in Krippe und Kindergarten - Konzept und Evaluation. In K. Fröhlich-Gildhoff, I. Nentwig-Gesemann & P. Strehmel (Hrsg.), *Forschung in der Frühpädagogik III. Schwerpunkt: Sprachentwicklung und Sprachförderung* (Materialien zur Frühpädagogik, Bd. 5, S. 107-133). Freiburg i. Br.: FEL Verl. Forschung Entwicklung Lehre.

- Butler, A. C., Godbole, N. & Marsh, E. J. (2013). Explanation feedback is better than correct answer feedback for promoting transfer of learning. *Journal of Educational Psychology*, 105 (2), 290-298.
- Butler, D. & Winne, P. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of educational research*, 65 (3), 245-281.
- Butler, R. (1987). Task-Involving and Ego-Involving Properties of Evaluation; Effects of Different Feedback Conditions on Motivational Perceptions, Interest, and Performance. *Journal of Educational Psychology*, 79 (4), 476-482.
- Cameron, H. K. (1966). The effectiveness of feedback in teaching principles of educational psychology. *Journal of Experimental Education*, 34 (3), 53-56.
- Cameron, J. & Pierce, W. D. (1994). Reinforcement, Reward, and Intrinsic Motivation: A Meta-Analysis. *Review of educational research*, 64 (3), 363-423.
- Carver, C. S. & Scheier, M. F. (1990). Origins and Functions of Positive and Negative Affect: A Control-Process View. *Psychological Review*, 97 (1), 19-35.
- Carver, C. S. & Scheier, M. F. (1981). *Attention and self-regulation. A control-theory approach to human behavior* (Springer series in social psychology). New York, NY: Springer.
- Chien, N. C., Howes, C., Burchinal, M., Pianta, R. C., Ritchie, S., Bryant, D. M. et al. (2010). Children's Classroom Engagement and School Readiness Gains in Prekindergarten. *Child Development*, 81 (5), 1534-1549.
- Cianci, A. M., Schaubroeck, J. M. & McGill, G. A. (2010). Achievement Goals, Feedback, and Task Performance. *Human Performance*, 23 (2), 131-154.
- Claiborn, C. D. & Goodyear, R. K. (2005). Feedback in psychotherapy. *Journal of clinical psychology*, 61 (2), 209-217.
- Clarke-Stewart, K. (1987). Predicting child development from child care forms and features. In D. Phillips (Hrsg.), *Quality in child care: what does research tell us?* (Research monographs of the National Association for the Education of Young Children, Vol. 1, S. 21-42). Washington, D.C: National Association for the Education of Young Children.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Orlando: Academic Press.
- Connor, C. M., Morrison, F. J. & Slominski, L. (2006). Preschool instruction and children's emergent literacy growth. *Journal of Educational Psychology*, 98 (4), 665-689.
- Crafts, L. W. & Gilbert, R. W. (1935). The effect of knowledge of results on maze learning and retention. *Journal of Educational Psychology*, 26 (3), 177-187.

- Curby, T. W., Rimm-Kaufman, S. E. & Ponitz, C. C. (2009). Teacher-child interactions and children's achievement trajectories across kindergarten and first grade. *Journal of Educational Psychology*, 101 (4), 912-925.
- Cutrim Schmid, E. (2011). Video-stimulated reflection as a professional development tool in interactive whiteboard research. *ReCALL*, 23 (03), 252-270.
- Dahling, J. J. & Ruppel, C. L. (2016). Learning goal orientation buffers the effects of negative normative feedback on test self-efficacy and reattempt interest. *Learning and Individual Differences*, 50, 296-301.
- De Rivera, C., Girolametto, L., Greenberg, J. & Weitzman, E. (2005). Children's Responses to Educators' Questions in Day Care Play Groups. *American Journal of Speech-Language Pathology*, 14(1), 14-26.
- Deci, E. L., Koestner, R. & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125 (6), 627-668.
- Deci, E. L. & Ryan, R. M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik*, 39(2), 223-238.
- Dennis, L. & Horn, E. (2013). The effects of professional development on preschool teachers' instructional behaviours during storybook reading. *Early Child Development and Care*, 184(8), 1160-1177.
- Deputy, E. C. (1929). Knowledge of Success as a Motivating Influence in College Work. *The Journal of Educational Research*, 20 (5), 327-334.
- Deutsche Liga für das Kind. (2008). *Gute Qualität in Krippe und Tagespflege. Positionspapier der Deutschen Liga für das Kind*. Berlin: Deutsche Liga für das Kind.
- Deutscher Bildungsrat. (1970). *Strukturplan für das Bildungswesen. Empfehlungen der Bildungskommission*. Stuttgart: Ernst Klett Verlag.
- Deutsches Jugendinstitut. (2008). *Stellungnahme des Deutschen Jugendinstitutes zum Entwurf eines Gesetzes zur Förderung von Kindern unter drei Jahren in Tageseinrichtungen und in der Kindertagespflege*. München: Deutsches Jugendinstitut e.V.
- Dickinson, D. K. & Tabors, P. O. (2001). *Beginning literacy with language. Young children learning at home and school*. Baltimore: P.H. Brookes Pub. Co.
- Döring, N. & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Dudenredaktion. (o.J.). "Feedback" auf Duden online. Zugriff am 25.05.2017. Verfügbar unter <http://www.duden.de/node/714633/revisions/1382751/view>
- Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, 41 (10), 1040-1048.

- Early, D. M., Maxwell, K. L., Burchinal, M., Alva, S., Bender, R. H., Bryant, D. et al. (2007). Teachers' education, classroom quality, and young children's academic skills: results from seven studies of preschool programs. *Child Development*, 78(2), 558-580.
- Egert, F. (2015). *Meta-analysis on the impact of in-service professional development programs for preschool teachers on quality ratings and child outcomes*. Dissertation. Otto-Friedrich-Universität, Bamberg.
- Elliot, A. J. & McGregor, H. A. (2001). A 2 × 2 achievement goal framework. *Journal of Personality and Social Psychology*, 80(3), 501-519.
- Elliott, A. J. & Murayama, K. (2008). On the Measurement of Achievement Goals: Critique Illustration, and Application. *Journal of Educational Psychology*, 100(3), 613-628.
- Elliott, E. S. & Dweck, C. S. (1988). Goals. An approach to motivation and achievement. *Journal of Personality and Social Psychology*, 54(1), 5-12.
- Elwell, J. L. & Grindley, G. C. (1938). The effect of knowledge of results on learning and performance. I. A co-ordinated movement of the two hands. *British Journal of Psychology*, 29(1), 39-53.
- Englund, L. W. (2010). *Evaluating and improving the quality of teacher's language modeling in early childhood classrooms*. Dissertation. University of Nevada, Las Vegas, NV.
- Ewing, A. R. & Taylor, A. R. (2009). The role of child gender and ethnicity in teacher-child relationship quality and children's behavioral adjustment in preschool. *Early Childhood Research Quarterly*, 24(1), 92-105.
- Fengler, J. (2010). Feedback als Interventions-Methode. *Gruppendynamik und Organisationsberatung*, 41(1), 5-20.
- Feys, M., Anseel, F. & Wille, B. (2011). Improving Feedback Reports. The Role of Procedural Information and Information Specificity. *Academy of Management Learning & Education*, 10(4), 661-681.
- Fried, L. (2011). Sprachförderstrategien in Kindergartengruppen – Einschätzungen und Ergebnisse mit DO-RESI. *Empirische Pädagogik*, 25(4), 543-562.
- Fried, L. & Briedigkeit, E. (2008). *Sprachförderkompetenz. Selbst- und Teamqualifizierung für Erzieherinnen, Fachberatungen und Ausbilder* (Frühe Kindheit Sprache & Literacy, 1. Aufl.). Berlin: Cornelsen Scriptor.
- Fröhlich-Gildhoff, K. (2014, Januar). „Was bedeutet Kompetenzorientierung in der Praxis?“. Vortrag am Staatsinstitut für Frühpädagogik, München.
- Fröhlich-Gildhoff, K., Nentwig-Gesemann, I. & Pietsch, S. (2011). *Kompetenzorientierung in der Qualifizierung frühpädagogischer Fachkräfte*. München: Deutsches Jugendinstitut e.V.
- Fthenakis, W. (2009). Ko-Konstruktion: Lernen durch Zusammenarbeit. *Kinderzeit*, 3, 8-13.

- Fukkink, R. & Lont, A. (2007). Does training matter? A meta-analysis and review of caregiver training studies. *Early Childhood Research Quarterly*, 22, 294-311.
- Fukkink, R. G. & Tavecchio, L. W. (2010). Effects of Video Interaction Guidance on early childhood teachers. *Teaching and Teacher Education*, 26(8), 1652-1659.
- Gasteiger-Klicpera, B., Knapp, W. & Kucharz, D. (2010). *Abschlussbericht der Wissenschaftlichen Begleitung des Programms "Sag' mal was – Sprachförderung für Vorschulkinder"*. Pädagogische Hochschule Weingarten. Verfügbar unter [http://www.ph-weingarten.de/zep/Projekte/Abschlussbericht\\_Sprachfoerderung\\_Landesstiftung\\_PH\\_Weingarten.pdf](http://www.ph-weingarten.de/zep/Projekte/Abschlussbericht_Sprachfoerderung_Landesstiftung_PH_Weingarten.pdf)
- Girolametto, L., Weitzmann, E. & Greenberg, J. (2000). *Teacher Interaction and Language Rating Scale*. Toronto: The Hanen Centre.
- Girolametto, L., Weitzmann, E. & Greenberg, J. (2003). Training Day Care Staff to Facilitate Children's Language. *American Journal of Speech-Language Pathology*, 12, 299-311.
- Girolametto, L. & Weitzman, E. (2002). Responsiveness of Child Care Providers in Interactions With Toddlers and Preschoolers. *Language Speech and Hearing Services in Schools*, 33(4), 268-282.
- Girolametto, L., Weitzman, E., Lefebvre, P. & Greenberg, J. (2007). The Effects of In-Service Education to Promote Emergent Literacy in Child Care Centers. A Feasibility Study. *LANGUAGE, SPEECH, AND HEARING SERVICES IN SCHOOLS*, 38, 72-83.
- Glüer, M. (2013). *Beziehungsqualität und kindliche Kooperations- und Bildungsbereitschaft. Eine Studie in Kindergarten und Grundschule* (Research). Wiesbaden: Springer VS.
- Göncü, A. & Weber, E. (2000). Preschoolers' Classroom Activities and Interactions with Peers and Teachers. *Early Education & Development*, 11(1), 93-107.
- Gordon, R. A., Fujimoto, K., Kaestner, R., Korenman, S. & Abner, K. (2013). An assessment of the validity of the ECERS-R with implications for measures of child care quality and relations to child development. *Developmental Psychology*, 49(1), 146-160.
- Grohmann, A. & Kauffeld, S. (2013). Evaluating training programs: development and correlates of the Questionnaire for Professional Training Evaluation. *International Journal of Training and Development*, 17(2), 135-155.
- Grossmann, K. & Grossmann, K. E. (2014). *Bindungen. Das Gefüge psychischer Sicherheit* (6. Aufl.). Stuttgart: Klett-Cotta.
- Hamre, B. K., Pianta, R. C., Burchinal, M., Field, S., LoCasale-Crouch, J., Downer, J. T. et al. (2012). A Course on Effective Teacher-Child Interactions. Effects on Teacher Beliefs, Knowledge, and Observed Practice. *American Educational Research Journal*, 49(1), 88-123.

- Hamre, B. K. & Pianta, R. C. (2005). Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure? *Child Development*, 76(5), 949-967.
- Hamre, B. K., Pianta, R. C., Downer, J. T., DeCoster, J., Mashburn, A. J., Jones, S. M. et al. (2013). Teaching through Interactions: Testing a developmental framework of teacher effectiveness in over 4,000 classrooms. *The Elementary School Journal*, 113(4), 461-487.
- Harms, T., Clifford, R. M. & Cryer, D. (1998). *Early childhood environment rating scale - Revised Edition (ECERS-R)*. New York: Teachers College Press.
- Hartley, H. O. (1950). The Maximum F-Ratio as a Short-Cut Test for Heterogeneity of Variance. *Biometrika*, 37(3/4), 308-312.
- Hattie, J. (1999). *Influences on student learning* (University of Auckland, Hrsg.). Zugriff am 16.07.2017. Verfügbar unter <https://cdn.auckland.ac.nz/assets/education/hattie/docs/influences-on-student-learning.pdf>
- Hattie, J. & Gan, M. (2011). Instruction based on feedback. In P. Alexander & R. Mayer (Hrsg.), *Handbook of research on learning and instruction* (S. 249-271). New York: Routledge.
- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of educational research*, 77(1), 81-112.
- Hattie, J. & Wollenschläger, M. (2014). A conceptualization of feedback. In H. Ditton & A. Müller (Hrsg.), *Feedback und Rückmeldungen. Theoretische Grundlagen, empirische Befunde, praktische Anwendungsfelder* (S. 135-149). Münster: Waxmann.
- Hedges, L. V. & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando: Academic Press.
- Helmerhorst, K. O. W., Riksen-Walraven, J. M. A., Fukkink, R. G., Tavecchio, L. W. C. & Gevers Deynoot-Schaub, M. J. J. M. (2017). Effects of the Caregiver Interaction Profile Training on Caregiver-Child Interactions in Dutch Child Care Centers. A Randomized Controlled Trial. *Child & Youth Care Forum*, 46(3), 413-436.
- Helmerhorst, K. O. W., Riksen-Walraven, J. M., Vermeer, H. J., Fukkink, R. G. & Tavecchio, L. W. C. (2014). Measuring the Interactive Skills of Caregivers in Child Care Centers. Development and Validation of the Caregiver Interaction Profile Scales. *Early Education & Development*, 25(5), 770-790.
- Helmke, A., Hornstein, W. & Terhart, E. (2000). Qualität und Qualitätssicherung im Bildungsbereich. Schule, Sozialpädagogik, Hochschule. *Zeitschrift für Pädagogik*, 41 (Beiheft), 7-16.



- Hensel, R., Meijers, F., van der Leeden, R. & Kessels, J. (2010). 360 degree feedback. How many raters are needed for reliable ratings on the capacity to develop competences, with personal qualities as developmental goals? *The International Journal of Human Resource Management*, 21 (15), 2813-2830.
- Heslin, P. A. & Latham, G. P. (2004). The Effect of Upward Feedback on Managerial Behavior. *Applied Psychology*, 53(1), 23-37.
- Hetherington, E. M., Parke, R. D. & Schmuckler, M. A. (2005). *Child psychology. A contemporary viewpoint* (2nd Canadian ed.). Toronto: McGraw-Hill Ryerson.
- Hildebrandt, F., Scheidt, A., Hildebrandt, A., Hédervári-Heller, É. & Dreier, A. (2016). Sustained shared thinking als Interaktionsformat und das Sprachverhalten von Kindern. *Frühe Bildung*, 5(2), 82-90.
- Hippel, A. v. & Grimm, R. (2010). *Qualitätsentwicklungskonzepte in der Weiterbildung Frühpädagogischer Fachkräfte. Expertise für das Projekt Weiterbildungsinitiative Frühpädagogische Fachkräfte (WiFF)*. München: Deutsches Jugendinstitut e.V. (DJI).
- Hofer, M., Wild, E. & Noack, P. (2002). *Lehrbuch Familienbeziehungen. Eltern und Kinder in der Entwicklung* (2., vollst. überarb. und erw. Aufl.). Göttingen u.a.: Hogrefe Verl. für Psychologie.
- Hoska, D. M. (1993). Motivating learners through CBI feedback: Developing a positive learner perspective. In J. Dempsey & G. Sales (Hrsg.), *Interactive instruction and feedback* (S. 105-132). Englewood Cliffs: Educational Technology Publications.
- Howes, C., Hamilton, C. E. & Matheson, C. C. (1994). Children's Relationships with Peers: Differential Associations with Aspects of the Teacher-Child Relationship. *Child Development*, 65, 253-263.
- Howes, C., Phillips, D. A. & Whitebook, M. (1992). Thresholds of Quality: Implications for the Social Development of Children in Center-Based Child Care. *Child Development*, 63(2), 449-460.
- Howes, C. & Stewart, P. (1987). Child's Play With Adults, Toys, and Peers: An Examination of Family and Child-Care Influences. *Developmental Psychology*, 23(3), 423-430.
- Howes, C., Burchinal, M., Pianta, R., Bryant, D., Early, D., Clifford, R. et al. (2008). Ready to learn? Children's pre-academic achievement in pre-Kindergarten programs. *Early Childhood Research Quarterly*, 23(1), 27-50.
- Howes, C., Matheson, C. C. & Hamilton, C. E. (1994). Maternal, Teacher, and Child Care History Correlates of Children's Relationships with Peers. *Child Development*, 65(1), 264-273.

- Howes, C. & Smith, E. W. (1995). Relations among child care quality, teacher behavior, children's play activities, emotional security, and cognitive activity in child care. *Early Childhood Research Quarterly*, 10(4), 381-404.
- Hundal, P. S. (1969). Knowledge of performance as an incentive in repetitive industrial work. *Journal of Applied Psychology*, 53(3, Pt.1), 224-226.
- Jacobson, N. S. & Truax, P. (1991). Clinical significance. A statistical approach to defining meaningful change in psychotherapy research. *Journal of consulting and clinical psychology*, 59(1), 12-19.
- Jugendministerkonferenz/Kultusministerkonferenz. (2004). *Gemeinsamer Rahmen der Länder für die frühe Bildung in Kindertageseinrichtungen*. Zugriff am 08.08.2017. Verfügbar unter [http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen\\_beschluesse/2004/2004\\_06\\_04-Fruehe-Bildung-Kitas.pdf](http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2004/2004_06_04-Fruehe-Bildung-Kitas.pdf)
- Justice, L., Mashburn, A., Hamre, B. & Pianta, R. (2008). Quality of language and literacy instruction in preschool classrooms serving at-risk pupils. *Early Childhood Research Quarterly*, 23(1), 51-68.
- Kammermeyer, G., Roux, S. & Stuck, A. (2013). *Was wirkt wie? Evaluation von Sprachfördermaßnahmen in Rheinland-Pfalz. Abschlussbericht*. Landau: Universität Koblenz /Landau. Zugriff am 08.08.2017.
- Kappauf, N. (2018). *Fachlich fit mit Videofeedback. Sprachliches Interaktionsverhalten in der Kita reflektieren*. (unter Mitarbeit von S. Fischer & C. Wirts). München: Staatsinstitut für Frühpädagogik.
- Karl, K. A., O'Leary-Kelly, A. M. & Martocchio, J. J. (1993). The Impact of Feedback and Self-Efficacy on Performance in Training. *Journal of Organizational Behavior*, 14(4), 379-394.
- Katz, L. (1996). Qualität der Früherziehung in Betreuungseinrichtungen: Fünf Perspektiven. In W. Tietze (Hrsg.), *Früherziehung. Trends, internationale Forschungsergebnisse, Praxisorientierungen* (S. 226-239). Neuwied: Luchterhand.
- Kaufman, D., Coddington, R. S., Markus, K. A., Tryon, G. S. & Kyse, E. N. (2013). Effects of verbal and written performance feedback on treatment adherence. Practical application of two delivery formats. *Journal of Educational & Psychological Consultation*, 23(4), 264-299.
- Kieferle, C. & Schauland, N. (2016). Video-Interaktionsbegleitung - eine Methode zur Optimierung der Fachkraft-Kind-Interaktion. *Frühe Bildung*, 5(2), 110-111.
- Kirkpatrick, D. (1998). *Evaluating Training Programs. The Four Levels*. San Francisco: Berrett-Koehler.

- Klein, S. M., Kraut, A. I. & Wolfson, A. (1971). Employee Reactions to Attitude Survey Feedback. A Study of the Impact of Structure and Process. *Administrative Science Quarterly*, 16 (4), 497-514.
- Kluger, A. & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119 (2), 254-284.
- König, A. (2009). *Interaktionsprozesse zwischen Erzieherinnen und Kindern. Eine Videostudie aus dem Kindergartenalltag* (1. Auflage). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Kontos, S. & Wilcox-Herzog, A. (1997). Influences on children's competence in early childhood classrooms. *Early Childhood Research Quarterly*, 12 (3), 247-262.
- Kopp, B. & Mandl, H. (2014). Lerntheoretische Grundlagen von Rückmeldungen. In H. Dittton & A. Müller (Hrsg.), *Feedback und Rückmeldungen. Theoretische Grundlagen, empirische Befunde, praktische Anwendungsfelder* (S. 29-41). Münster: Waxmann.
- Krampen, G. (1987). Differential effects of teacher comments. *Journal of Educational Psychology*, 79 (2), 137-146.
- Krause, U.-M. (2007). *Feedback und kooperatives Lernen* (Pädagogische Psychologie und Entwicklungspsychologie, Bd. 60). Zugl.: München, Univ., Diss., 2005. Münster: Waxmann.
- Krumboltz, J. D. & Weisman, R. G. (1962). The effect of intermittent confirmation in programmed instruction. *Journal of Educational Psychology*, 53 (6), 250-253.
- Kucharz, D., Mackowiak, K. & Beckerle, C. (2015). *Alltagsintegrierte Sprachförderung. Ein Konzept zur Weiterqualifizierung in Kita und Grundschule*. Weinheim: Beltz Verlag.
- Kulhavy, R. (1977). Feedback in written instruction. *Journal of Educational Research*, 47 (1), 211-232.
- Kulhavy, R. & Stock, W. (1989). Feedback in written instruction: The place of response certificate. *Educational Psychology Review*, 1 (4), 279-308.
- Kulhavy, R. & Wager, W. (1993). Feedback in programmed instruction: Historical context and implications for practice. In J. V. Dempsey & G. C. Sales (Hrsg.), *Interactive instruction and feedback* (1. print, S. 3-20). Englewood Cliffs NJ: Educational Technology Publ.
- Kulhavy, R. W. & Anderson, R. C. (1972). Delay-retention effect with multiple-choice tests. *Journal of Educational Psychology*, 63 (5), 505-512.
- Kulhavy, R. W., White, M. T., Topp, B. W., Chan, A. L. & Adams, J. (1985). Feedback complexity and corrective efficiency. *Contemporary Educational Psychology*, 10 (3), 285-291.
- Kulik, J. & Kulik, C.-L. (1988). Timing of feedback and verbal learning. *Review of educational research*, 58 (1), 79-97.

- Ladd, G. W. & Burgess, K. B. (2001). Do Relational Risks and Protective Factors Moderate the Linkages between Childhood Aggression and Early Psychological and School Adjustment? *Child Development*, 72(5), 1579-1601.
- Lerkkanen, M.-K., Kiuru, N., Pakarinen, E., Poikkeus, A.-M., Rasku-Puttonen, H., Siekkinen, M. et al. (2016). Child-centered versus teacher-directed teaching practices. Associations with the development of academic skills in the first grade at school. *Early Childhood Research Quarterly*, 36, 145-156.
- Linderbaum, B. & Levy, P. (2010). The development and validation of the feedback orientation scale (FOS). *Journal of Management*, 36(6), 1372-1405.
- Locke, E. A. & Latham, G. P. (1990). *A theory of goal setting and task performance*. Englewood Cliffs: Prentice-Hall.
- Locke, E. A., Shaw, K. N., Saari, L. M. & Latham, G. P. (1981). Goal setting and task performance. 1969-1980. *Psychological Bulletin*, 90(1), 125-152.
- Lublin, S. C. (1965). Reinforcement schedules, scholastic aptitude, autonomy need, and achievement in a programmed course. *Journal of Educational Psychology*, 56(6), 295-302.
- Lyden, J. A., Chaney, L. H., Daneshmandi, V. & Houston, D. A. (2002). Anchoring, Attributions, and Self-Efficacy. An Examination of Interactions. *Contemporary Educational Psychology*, 27(1), 99-117.
- Mähler, C. & Stern, E. (2006). Transfer. In D. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie*. (S. 782-793). Weinheim: Beltz Verlag.
- Main, M. & Solomon, J. (1986). Discovery of an Insecure-Disorganized/Disoriented Attachment Pattern. In T. B. Brazelton & M. W. Yogman (Hrsg.), *Affective development in infancy* (S. 95-124). Norwood, NJ: Ablex Publ. Corp.
- Malone, T. W. (1981). Toward a Theory of Intrinsically Motivating Instruction. *Cognitive Science*, 5(4), 333-369.
- Markussen-Brown, J., Juhl, C. B., Piasta, S. B., Bleses, D., Højen, A. & Justice, L. M. (2017). The effects of language- and literacy-focused professional development on early educators and children. A best-evidence meta-analysis. *Early Childhood Research Quarterly*, 38, 97-115.
- Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D. et al. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development*, 79(3), 732-749.
- Mason, B. J. & Bruning, R. H. (2001). *Providing feedback in computer-based instruction: What the research tells us. CLASS Project Research Report No. 9* (Center for Instructional Innovation, Hrsg.). University of Nebraska-Lincoln.

- Massey, S. L., Pence, K. L., Justice, L. M. & Bowles, R. P. (2008). Educators' Use of Cognitively Challenging Questions in Economically Disadvantaged Preschool Classroom Contexts. *Early Education & Development*, 19(2), 340-360.
- Mayer, D. & Beckh, K. (2016). Examining the validity of the ECERS-R. Results from the German National Study of Child Care in Early Childhood. *Early Childhood Research Quarterly*, 36, 415-426.
- Mayer, D. & Beckh, K. (2018). Erfassung pädagogischer Qualität in Kindertageseinrichtungen. *Frühe Bildung*, 7(2), 67-76.
- Mayer, D., Beckh, K., Berkic, J. & Becker-Stoll, F. (2013). *Nationale Untersuchung zur Bildung, Betreuung und Erziehung in der frühen Kindheit (NUBBEK). Ergebnisbericht der Untersuchung in Bayern*. München: Staatsinstitut für Frühpädagogik.
- McCartney, K. (1984). Effect of quality of day care environment on children's language development. *Developmental Psychology*, 20(2), 244-260.
- McDonald Connor, C., Son, S.-H., Hindman, A. H. & Morrison, F. J. (2005). Teacher qualifications, classroom practices, family characteristics, and preschool experience. Complex effects on first graders' vocabulary and early reading outcomes. *Journal of School Psychology*, 43(4), 343-375.
- Milburn, T. F., Girolametto, L., Weitzman, E. & Greenberg, J. (2014). Enhancing preschool educators' ability to facilitate conversations during shared book reading. *Journal of Early Childhood Literacy*, 14(1), 105-140.
- Miller, G. A., Galanter, E. & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York NY: Holt Rinehart and Winston.
- Mory, E. (2004). Feedback research revisited. In D. Jonassen (Hrsg.), *Handbook of research for educational communications and technology* (S. 745-783). Mahwah: Lawrence Erlbaum.
- Müller, A. & Ditton, H. (2014). Feedback: Begriff, Formen und Funktionen. In H. Ditton & A. Müller (Hrsg.), *Feedback und Rückmeldungen. Theoretische Grundlagen, empirische Befunde, praktische Anwendungsfelder* (S. 11-28). Münster: Waxmann.
- Murphy, K., Cleveland, J. & Mohler, C. (2001). Reliability, validity, and meaningfulness of multisource ratings. In D. Bracken, C. W. Timmreck & A. H. Church (Eds.), *The handbook of multisource feedback. The comprehensive resource for designing and implementing MSF processes* (The Jossey-Bass business & management series, 1st ed., pp. 130-148). San Francisco: Jossey-Bass.
- Narciss, S. (2014). Modelle zu den Bedingungen und Wirkungen von Feedback in Lehr-Lernsituationen. In H. Ditton & A. Müller (Hrsg.), *Feedback und Rückmeldungen. Theorie-*

- tische Grundlagen, empirische Befunde, praktische Anwendungsfelder* (S. 43–82). Münster: Waxmann.
- Narciss, S. (2004). The Impact of Informative Tutoring Feedback and Self-Efficacy on Motivation and Achievement in Concept Learning. *Experimental Psychology*, 51 (3), 214–228.
- Narciss, S. (2006). *Informatives tutorielles Feedback. Entwicklungs- und Evaluationsprinzipien auf der Basis instruktionspsychologischer Erkenntnisse* (Pädagogische Psychologie und Entwicklungspsychologie, Bd. 56). Zugl.: Dresden, Techn. Univ., Habil.-Schr., 2005. Münster: Waxmann.
- Nentwig-Gesemann, I., Fröhlich-Gildhoff, K. & Pietsch, S. (2011). Kompetenzentwicklung von FrühpädagogInnen in Aus- und Weiterbildung. *Frühe Bildung*, 22–30.
- NICHD Early Child Care Research Network. (2002). Early Child Care and Children's Development Prior to School Entry. Results from the NICHD Study of Early Child Care. *American Educational Research Journal*, 39 (1), 133–164.
- NICHD Early Child Care Research Network. (2003). Social Functioning in First Grade. Associations With Earlier Home and Child Care Predictors and With Current Classroom Experiences. *Child Development*, 74 (6), 1639–1662.
- NICHD ECCRN. (2002). Child-Care Structure → Process → Outcome. Direct and Indirect Effects of Child-Care Quality on Young Children's Development. *Psychological Science*, 13 (3), 199–206.
- Pakarinen, E., Lerkkanen, M.-K., Poikkeus, A.-M., Kiuru, N., Siekkinen, M., Rasku-Puttonen, H. et al. (2010). A Validation of the Classroom Assessment Scoring System in Finnish Kindergartens. *Early Education and Development*, 21 (1), 95–124.
- Payne, S. C., Youngcourt, S. S. & Beaubien, J. M. (2007). A meta-analytic examination of the goal orientation nomological net. *Journal of Applied Psychology*, 92 (1), 128–150.
- Peisner-Feinberg, E. S., Burchinal, M. R., Clifford, R. M., Culkin, M. L., Howes, C., Kagan, S. L. et al. (2001). The Relation of Preschool Child-Care Quality to Children's Cognitive and Social Developmental Trajectories through Second Grade. *Child Development*, 72 (5), 1534–1553.
- Pentimonti, J. M. & Justice, L. M. (2010). Teachers' Use of Scaffolding Strategies During Read Alouds in the Preschool Classroom. *Early Childhood Education Journal*, 37 (4), 241–248.
- Pentimonti, J. M., Justice, L. M., Yeomans-Maldonado, G., McGinty, A. S., Slocum, L. & O'Connell, A. (2017). Teachers' Use of High- and Low-Support Scaffolding Strategies to Differentiate Language Instruction in High-Risk/Economically Disadvantaged Settings. *Journal of Early Intervention*, 39 (2), 125–146.

- Piaget, J. (1992). *Psychologie der Intelligenz* (3., in der Ausstattung veränd. Aufl.). Stuttgart: Klett-Cotta.
- Pianta, R. C. (2017). Beobachtung und Weiterentwicklung der Fachkraft-Kind-Interaktionen in der Frühpädagogik. In M. Wertfein, A. Wildgruber, C. Wirts & F. Becker-Stoll (Hrsg.), *Interaktionen in Kindertageseinrichtungen. Theorie und Praxis im interdisziplinären Dialog* (S. 22-34). Göttingen: Vandenhoeck & Ruprecht.
- Pianta, R. C., La Paro, K. M. & Hamre, B. K. (2008). *Classroom Assessment Scoring System. Manual Pre-K* (7. Aufl.). CLASS Pre-K. Baltimore: Paul Brookes.
- Pianta, R. C., Mashburn, A. J., Downer, J. T., Hamre, B. K. & Justice, L. (2008). Effects of web-mediated professional development resources on teacher-child interactions in pre-kindergarten classrooms. *Early Childhood Research Quarterly* (23), 431-451.
- Pianta, R. C. & Stuhlman, M. W. (2004). Teacher-Child Relationships and Children's Success in the First Years of School. *School Psychology Review*, 33 (3), 444-458.
- Pianta, R. (2001). *Student-teacher relationship scale. Professional manual*. Lutz: Psychological Assessment Resources.
- Pianta, R. C., Belsky, J., Houts, R. & Morrison, F. (2007). Teaching. Opportunities to learn in America's elementary classrooms. *Science (New York, N.Y.)*, 315 (5820), 1795-1796.
- Pianta, R. C., La Paro, K. M., Payne, C., Cox, M. J. & Bradley, R. (2002). The Relation of Kindergarten Classroom Environment to Teacher, Family, and School Characteristics and Child Outcomes. *The Elementary School Journal*, 102 (3), 225-238.
- Powell, D. R., Diamond, K. E., Burchinal, M. R. & Koehler, M. J. (2010). Effects of an early literacy professional development intervention on head start teachers and children. *Journal of Educational Psychology*, 102 (2), 299-312.
- Pramling Samuelsson, I. (2009). Lernen und Lernprozesse in der Frühpädagogik. In F. Becker-Stoll (Hrsg.), *Bildung und Erziehung in Deutschland. Pädagogik für Kinder von 0 bis 10 Jahren* (Frühe Kindheit : Pädagogische Ansätze, 1. Aufl., S. 28-45). Berlin u.a.: Cornelsen Scriptor.
- Radan, J., Schauland, N. & Wirts, C. (2017). *BiSS-Analyseschema zur sprachlichen Interaktion - Video und Transkripte*. Unveröffentlichtes Manuskript, IFP: München.
- Raemdonck, I. & Strijbos, J.-W. (2013). Feedback perceptions and attribution by secretarial employees. Effects of feedback-content and sender characteristics. *European Journal of Training and Development*, 37 (1), 24-48.
- Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science*, 28 (1), 4-13.
- Reichert-Garschhammer, E. & Kieferle, C. (2011). *Sprachliche Bildung in Kindertageseinrichtungen*. Freiburg: Herder Verlag.

- Relich, J. D., Debus, R. L. & Walker, R. (1986). The mediating role of attribution and self-efficacy variables for treatment effects on achievement outcomes. *Contemporary Educational Psychology*, 11 (3), 195-216.
- Remsperger, R. (2011). *Sensitive Responsivität. Zur Qualität pädagogischen Handelns im Kindergarten* (1. Auflage). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Rimm-Kaufman, S. E., Curby, T. W., Grimm, K. J., Nathanson, L. & Brock, L. L. (2009). The contribution of children's self-regulation and classroom quality to children's adaptive behaviors in the kindergarten classroom. *Developmental Psychology*, 45 (4), 958-972.
- Rimm-Kaufman, S. E., Early, D. M., Cox, M. J., Saluja, G., Pianta, R. C., Bradley, R. H. et al. (2002). Early behavioral attributes and teachers' sensitivity as predictors of competent behavior in the kindergarten classroom. *Journal Of Applied Developmental Psychology*, 23 (4), 451-470.
- Salomon, G. & Globerson, T. (1987). Skill may not be enough: The role of mindfulness in learning and transfer. *International Journal of Educational Research*, 11 (6), 623-637.
- Sansone, C. (1986). A question of competence. The effects of competence and task feedback on intrinsic interest. *Journal of Personality and Social Psychology*, 51 (5), 918-931.
- Sassenrath, J. M. (1975). Theory and results on feedback and retention. *Journal of Educational Psychology*, 67 (6), 894-899.
- Schiefele, U. & Schreyer, I. (1994). Intrinsische Lernmotivation und Lernen: ein Überblick zu Ergebnissen der Forschung. *Zeitschrift für Pädagogische Psychologie*, 8 (1), 1-13.
- Schmidt, B. (2010). Qualitätsentwicklung = Evaluation + Hochschuldidaktik? Von einer Rechnung, die nicht aufgeht. In P. Pohlenz & A. Oppermann (Hrsg.), *Lehre und Studium professionell evaluieren. Wie viel Wissenschaft braucht die Evaluation?* (S. 179-192). Bielefeld: Universitätsverlag Weblar.
- Schröder, L. & Keller, H. (2013). *Sprachliche Bildung im Kita-Alltag. Gespräche mit Kindern anregen und lebendig gestalten*. (Begleitheft zum Film). Berlin: Cornelsen.
- Schuler, H., Prochaska, M. & Frintrup, A. (2001). *Leistungsmotivationsinventar. Dimensionen berufsbezogener Leistungsorientierung*. Göttingen: Hogrefe.
- Schulz von Thun, F. (2013). *Miteinander reden: Störungen und Klärungen. Allgemeine Psychologie der Kommunikation*. Hamburg: Rowohlt Verlag.
- Schunk, D. H. (1983). Ability Versus Effort Attributional Feedback: Differential Effects on Self-Efficacy and Achievement. *Journal of Educational Psychology*, 75 (6), 848-856.
- Schunk, D. H. & Swartz, C. W. (1993). Goals and Progress Feedback. Effects on Self-Efficacy and Writing Achievement. *Contemporary Educational Psychology*, 18 (3), 337-354.
- Shrout, P. E. & Fleiss, J. L. (1979). Intraclass Correlations: Uses in Assessing Rater Reliability. *Psychological Bulletin*, 86 (2), 420-428.



- Shute, V. (2008). Focus on formative feedback. *Review of educational research*, 78(1), 153–189.
- Siegler, R., DeLoache, J. & Eisenberg, N. (2008). *Entwicklungspsychologie im Kindes- und Jugendalter*. Heidelberg: Spektrum Akademischer Verlag.
- Simon, S. & Sachse, S. (2011). Sprachförderung in der Kindertagesstätte-Verbessert ein Interaktionstraining das sprachförderliche Verhalten von Erzieherinnen? *Empirische Pädagogik*, 25(4), 462–480.
- Simon, S. & Sachse, S. (2013). Anregung der Sprachentwicklung durch ein Interaktionstraining für Erzieherinnen. *Diskurs Kindheits- und Jugendforschung* (4), 379–397.
- Siraj-Blatchford, I. (2007). Effektive Bildungsprozesse: Lehren in der frühen Kindheit. In F. Becker-Stoll & M. Textor (Hrsg.), *Die Erzieherin-Kind-Beziehung. Zentrum von Bildung und Erziehung* (S. 98–114). Mannheim: Cornelsen Scriptor.
- Siraj-Blatchford, I., Sylva, K., Muttock, S., Gilden, R. & Bell, D. (2002). *Researching Effective Pedagogy in the Early Years*. Oxford: Department of Educational Studies.
- Siraj-Blatchford, I. & Manni, L. (2008). 'Would you like to tidy up now?'. An analysis of adult questioning in the English Foundation Stage. *Early Years*, 28(1), 5–22.
- Skinner, B. F. (1958). Teaching Machines. From the experimental study of learning come devices which arrange optimal conditions for self-instruction. *Science*, 128(3330), 969–977.
- Skinner, B. F. (1968). *The technology of teaching* (The century psychology series). Englewood Cliffs NJ: Prentice-Hall.
- Smith, M. W. & Dickinson, D. K. (2002). *Early language & literacy classroom observation*. Baltimore: Paul H. Brookes.
- Statistisches Bundesamt. (2016). *Betreuungsquoten der Kinder unter 6 Jahren in Kindertagesbetreuung am 01.03.2016*. Zugriff am 29.12.2017. Verfügbar unter [https://www.destatis.de/DE/ZahlenFakten/GesellschaftStaat/Soziales/Sozialleistungen/Kindertagesbetreuung/Tabellen/Tabellen\\_Betreuungsquote.html;jsessionid=DE11ECCF65B8AB25D18307E0702C7F56.InternetLive1](https://www.destatis.de/DE/ZahlenFakten/GesellschaftStaat/Soziales/Sozialleistungen/Kindertagesbetreuung/Tabellen/Tabellen_Betreuungsquote.html;jsessionid=DE11ECCF65B8AB25D18307E0702C7F56.InternetLive1)
- Strijbos, J.-W. & Müller, A. (2014). Personale Faktoren im Feedbackprozess. In H. Ditton & A. Müller (Hrsg.), *Feedback und Rückmeldungen. Theoretische Grundlagen, empirische Befunde, praktische Anwendungsfelder* (S. 83–134). Münster: Waxmann.
- Suchodoletz, A. von, Fäsche, A., Gunzenhauser, C. & Hamre, B. K. (2014). A typical morning in preschool. Observations of teacher-child interactions in German preschools. *Early Childhood Research Quarterly*, 29(4), 509–519.

- Sylva, K. (2017). Wirksame Lernunterstützung in der frühkindlichen Bildung und Betreuung. In M. Wertfein, A. Wildgruber, C. Wirts & F. Becker-Stoll (Hrsg.), *Interaktionen in Kindertageseinrichtungen. Theorie und Praxis im interdisziplinären Dialog* (S. 35-46). Göttingen: Vandenhoeck & Ruprecht.
- Sylva, K., Melhuish, E., Sammons, P., Blatchford, I. S. & Taggart, B. (2014). *The Effective Provision of Pre-school Education (EPPE): Final report*. Annesley: Department for Education and Skills.
- Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I. & Taggart, B. (2004). *The Effective Provision of Pre-school Education (EPPE) Project: Findings from pre-school to end of key stage 1*. Nottingham: Department for Education and Skills.
- Sylva, K., Siraj-Blatchford, I. & Taggart, B. (2011). *ECERS-E: The Four Curricular Subscales Extension to the Early Childhood Environment Rating Scale (ECERS-R)* (4. Aufl.). New York: Teachers College Press.
- Taylor, B. M., Pearson, P. D., Peterson, D. S. & Rodriguez, M. C. (2003). Reading Growth in High-Poverty Classrooms. The Influence of Teacher Practices That Encourage Cognitive Engagement in Literacy Learning. *The Elementary School Journal*, 104(1), 3-28.
- Teachstone. (2014). *Feedback Strategies. Sharing Observations with Teachers*. Participant Guide (Teachstone, Hrsg.). Charlottesville, Virginia.
- Tenorth, H.-E. & Tippelt, R. (Hrsg.). (2007). *Beltz Lexikon Pädagogik*. Weinheim: Beltz.
- Textor, M. (2007). Bildung in der Erzieherin-Kind-Interaktion. In F. Becker-Stoll & M. Textor (Hrsg.), *Die Erzieherin-Kind-Beziehung. Zentrum von Bildung und Erziehung* (S. 74-96). Mannheim: Cornelsen Scriptor.
- Thorndike, E. L. (1920). *Educational psychology. The original nature of man* (Bd. 1). New York: Teachers College, Columbia University.
- Tietze, W., Meischner, T., Gänsfuß, R., Grenner, K., Schuster, K.-M. & Völkel, P. & Roßbach, H.-G. (1998). *Wie gut sind unsere Kindergärten? Eine Untersuchung zur pädagogischen Qualität in deutschen Kindergärten*. Neuwied: Luchterhand.
- Tietze, W., Schuster, K.-M., Grenner, K. & Roßbach, H.-G. (2001). *Kindergarten-Skala, Revidierte Fassung (KES-R). (Deutsche Fassung der Early Childhood Environment Rating Scale Revised Edition von Thelma Harms, Richard M. Clifford, & Debby Cryer)*. Neuwied: Luchterhand.
- Tomasello, M. (2009). *Die Ursprünge der menschlichen Kommunikation* (Erste Auflage). Frankfurt am Main: Suhrkamp.
- Tures, A. (2014). *Im Dialog mit jungen Kindern. Einblicke in die Professionalisierungsprozesse von Frühpädagogikstudierenden ; eine interdisziplinäre und multimethodische*

- Studie* (International Cultural-historical Human Sciences, Bd. 50, neue Ausg). Zugl.: München, Univ., Diss., 2014. Berlin: Lehmanns.
- Vasilyeva, M., Huttenlocher, J. & Waterfall, H. (2006). Effects of language intervention on syntactic skill levels in preschoolers. *Developmental Psychology*, 42(1), 164-174.
- Wasik, B., Bond, M. & Hindman, A. (2006). The effects of a language and literacy intervention on head start children and teachers. *Journal of Educational Psychology*, 98(1), 63-74.
- Waters, E. (1995). Appendix A: The Attachment Q-Set (Version 3.0). *Monographs of the Society for Research in Child Development*, 60(2/3), 234-246.
- Waters, E. & Deane, K. E. (1985). Defining and Assessing Individual Differences in Attachment Relationships: Q-Methodology and the Organization of Behavior in Infancy and Early Childhood. In I. Bretherton & E. Waters (Hrsg.), *Growing Points of Attachment Theory and Research. Monographs of the Society for Research in Child Development* (Bd. 209, S. 41-65). Chicago: University of Chicago Press.
- Watzlawick, P., Bavelas, J. B. & Jackson, D. D. (2007). *Menschliche Kommunikation. Formen, Störungen, Paradoxien* (11., unveränderte Auflage). Bern: Verlag Hans Huber.
- Weiner, B. (1994). *Motivationspsychologie* (3. Aufl.). Weinheim: Beltz Psychologie Verl.-Union.
- Weinert, F. E. (Hrsg.). (2001). *Leistungsmessungen in Schulen* (Beltz-Pädagogik). Weinheim: Beltz.
- Weiterbildungsinitiative Frühpädagogische Fachkräfte. (2011). *Sprachliche Bildung. Grundlagen für die kompetenzorientierte Weiterbildung. Ein Wegweiser der Weiterbildungsinitiative Frühpädagogische Fachkräfte (WiFF)* (Nr. 1). München: Deutsches Jugendinstitut e.V. (DJI). Verfügbar unter <https://www.weiterbildungsinitiative.de/publikationen/details/data/sprachliche-bildung/?L=0>
- Weltzien, D. (2013). Erfassung von Interaktionsgelegenheiten im Alltag- erste Ergebnisse der Entwicklung und Überprüfung des Beobachtungsverfahrens GlnA. In K. Fröhlich-Gildhoff, I. Nentwig-Gesemann, A. König, U. Stenger & D. Weltzien (Hrsg.), *Forschung in der Frühpädagogik VI. Schwerpunkt: Interaktion zwischen Fachkräften und Kindern* (Materialien zur Frühpädagogik, Bd. 12, S. 59-85). Freiburg: FEL Verlag Forschung-Entwicklung-Lehre.
- Weltzien, D. (2014). *Pädagogik: Die Gestaltung von Interaktionen in der Kita. Merkmale-Beobachtung-Reflexion*. Weinheim und Basel: Beltz Juventa.
- Werner, C. D., Linting, M., Vermeer, H. J. & van IJzendoorn, M. H. (2016). Do Intervention Programs in Child Care Promote the Quality of Caregiver-Child Interactions? A Meta-Analysis of Randomized Controlled Trials. *Prevention Science*, 17(2), 259-273.

- Werner, C. D., Vermeer, H. J., Linting, M. & van IJzendoorn, M. H. (2018). Video-feedback intervention in center-based child care. A randomized controlled trial. *Early Childhood Research Quarterly*, 42, 93-104.
- Wharton-McDonald, R., Pressley, M. & Hampston, J. M. (1998). Literacy Instruction in Nine First-Grade Classrooms. Teacher Characteristics and Student Achievement. *The Elementary School Journal*, 99(2), 101-128.
- Whitehurst, G. J., Arnold, D. S., Epstein, J. N., Angell, A. L., Smith, M. & Fischel, J. E. (1994). A picture book reading intervention in day care and home for children from low-income families. *Developmental Psychology*, 30(5), 679-689.
- Wildgruber, A., Wirts, C. & Wertfein, M. (2014). Interaktionsqualität in Kindertageseinrichtungen in Deutschland – Forschung mit dem „Classroom Assessment Scoring System“. In A. Prengel & U. Winklhofer (Hrsg.), *Kinderrechte in pädagogischen Beziehungen. Band 2: Forschungszugänge* (S. 183-193). Opladen: Verlag Barbara Budrich.
- Winne, P. H., Graham, L. & Prock, L. (1993). A Model of Poor Readers' Text-Based Inferencing. Effects of Explanatory Feedback. *Reading Research Quarterly*, 28(1), 52-66.
- Wirts, C., Wertfein, M. & Wildgruber, A. (2017). Unterstützung kindlicher Kompetenzentwicklung und ihre Bedingungen in Kindertageseinrichtungen. In M. Wertfein, A. Wildgruber, C. Wirts & F. Becker-Stoll (Hrsg.), *Interaktionen in Kindertageseinrichtungen. Theorie und Praxis im interdisziplinären Dialog* (S. 59-72). Göttingen: Vandenhoeck & Ruprecht.
- Wirtz, M. & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität. Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen*. Göttingen: Hogrefe Verl. für Psychologie.
- Wissenschaftlicher Beirat für Familienfragen beim BMFSFJ. (2008). *Bildung, Betreuung und Erziehung für Kinder unter drei Jahren - elterliche und öffentliche Fürsorge in gemeinsamer Verantwortung* (Bundesministerium für Familie, Senioren, Frauen und Jugend, Hrsg.).
- Wollenschläger, M., Möller, J. & Harms, U. (2011). Effekte kompetenzieller Rückmeldung beim wissenschaftlichen Denken. *Zeitschrift für Pädagogische Psychologie*, 25(3), 197-202.
- Wygotski, L. S. (1969). *Denken und Sprechen*. Frankfurt am Main: Fischer.
- Wygotski, L. S. (1987). *Ausgewählte Schriften. Arbeiten zur psychischen Entwicklung der Persönlichkeit* (Bd. 2, 1. Aufl.). Berlin: Volk und Wissen.
- Zan, B. & Donegan-Ritter, M. (2014). Reflecting, coaching, and mentoring to enhance teacher-child interactions in head start classrooms. *Early Childhood Education Journal*, 42, 93-104.

# 11 Tabellenverzeichnis

Tabelle 1: Klassifikation unterschiedlicher Feedbackformen nach Narciss (2006, S. 23).....	40
Tabelle 2: Übersicht über postulierte Feedback-Funktionen nach Narciss (2006, S. 77).....	42
Tabelle 3: Implikationen für die Gestaltung von Feedback .....	69
Tabelle 4: Dimension "Language Modeling", entnommen aus "Classroom Assessment Scoring System Manual, PRE-K" (CLASS Pre-K; S. 75; Pianta, La Paro et al., 2008; Teachstone Training, LLC, Charlottesville, VA; Reprinted by permission).....	90
Tabelle 5: Kategorisierung von Fragetypen nach Siraj-Blatchford und Manni (2008).....	93
Tabelle 6: Kodierung des sprachlichen Interaktionsverhaltens.....	95
Tabelle 7: Skalen des „Questionnaire for Professional Training Evaluation“ von Grohmann und Kauffeld (2013) sowie die dazugehörigen Cronbachs Alpha-Werte.....	100
Tabelle 8: Erfassung personenbezogener Merkmale in der vorliegenden Studie.....	106
Tabelle 9: Items und Skalen der hier eingesetzten deutschen Übersetzung des AGQ-R sowie die dazugehörigen Cronbachs Alpha-Werte aus der vorliegenden Studie und der Studie von Elliott und Murayama (2008).....	108
Tabelle 10: Deskriptive Kennwerte der Skalen zur Zielorientierung – vergleichend für Feedbackgruppe (N = 44) und Vergleichsstichprobe (N = 229) von Elliott und Murayama (2008).....	109
Tabelle 11: Items der hier eingesetzten deutschen Übersetzung des FOS sowie die ermittelten Cronbachs Alpha-Werte für alle Skalen – vergleichend für die vorliegende Studie und die Studien von Linderbaum und Levy (2010).....	110
Tabelle 12: Deskriptive Kennwerte der Skalen aus der FOS – vergleichend für Feedbackgruppe (N = 44) und Vergleichsstichproben aus Studie 1 (N = 280) und Studie 2 (N = 267).....	111
Tabelle 13: Erfassung situationsspezifischer Merkmale in der vorliegenden Studie .....	112
Tabelle 14: Erfasste Variablen in der vorliegenden Studie .....	113
Tabelle 15: Deskriptive Statistik für Demografische Merkmale der Stichprobe.....	116
Tabelle 16: Deskriptive Statistik für den fachlichen/beruflichen Hintergrund der Fachkräfte ...	117
Tabelle 17: Beschreibung der Stichprobe auf Kita-Ebene .....	118
Tabelle 18: Häufigkeit von technischen Problemen während der Feedbackgespräche.....	132
Tabelle 19: In den Feedbackgesprächen thematisierte Indikatoren der Dimension „Language Modeling“ .....	134
Tabelle 20: Bewertung des Feedbackgesprächs durch die Fachkräfte: Deskriptive Statistik für die fünf Skalen des Q4TE .....	135
Tabelle 21: Deskriptive Statistik der Variable "Language Modeling" unterschieden nach Bedingung und Messzeitpunkt.....	138
Tabelle 22: Mittlere Anzahl an sprachlichen Verhaltensweisen während einer Bilderbuchbetrachtung für die Gesamtstichprobe zu T1 (N = 96).....	142

Tabelle 23: ICC-Koeffizienten für die einzelnen sprachlichen Verhaltensweisen .....	143
Tabelle 24: Deskriptive Statistik zu den beiden Haupteffekten der gestellten Fragen .....	145
Tabelle 25: Deskriptive Statistik der gestellten Fragen – unterschieden nach Bedingung und Messzeitpunkt .....	146
Tabelle 26: Deskriptive Kennwerte für die Aufforderungen zum Sprechen mit und ohne Antwortmöglichkeit .....	150
Tabelle 27: Deskriptive Statistik für die Variable "Anknüpfen" .....	153
Tabelle 28: Deskriptive Kennwerte für die Anzahl der längeren Gespräche .....	154
Tabelle 29: Deskriptive Kennwerte für das handlungsbegleitende Sprechen .....	155
Tabelle 30: Deskriptive Statistik für die Haupteffekte in Bezug auf die Wiederholung und Erweiterung kindlicher Äußerungen .....	156
Tabelle 31: Deskriptive Kennwerte für die Wiederholung und Erweiterung kindlicher Äußerungen – unterschieden nach Bedingung und Messzeitpunkt .....	157
Tabelle 32: Inferenzstatistische Kennwerte für den Interaktionseffekt – Wiederholung und Erweiterung .....	157
Tabelle 33: Signifikanz und Effektstärken der berechneten Interaktionseffekte für die Anzahl sprachlicher Verhaltensweisen .....	161
Tabelle 34: Häufigkeitsangaben zu Veränderungen nach dem "Reliable Change Index" von Jacobson und Truax (1991) .....	163
Tabelle 35: Deskriptive Kennwerte zu den situativen Rahmenbedingungen bei der Bilderbuchbetrachtung .....	166
Tabelle 36: Zielvereinbarungen .....	230
Tabelle 37: Häufigkeit der verschiedenen Frageformen zu T1 .....	231
Tabelle 38: Kennwerte für den Haupteffekt der Bedingung in Bezug auf das sprachliche Verhalten .....	232
Tabelle 39: Kennwerte für den Haupteffekt des Messzeitpunkts in Bezug auf das sprachliche Verhalten .....	233
Tabelle 40: Kennwerte der Chi-Quadrat-Tests zur Ermittlung von Unterschieden in kategorialen Störvariablen .....	234
Tabelle 41: Kennwerte der t-Tests für unabhängige Stichproben zur Ermittlung von Unterschieden in metrischen Störvariablen .....	235
Tabelle 42: Kennwerte der Interaktion zwischen Bedingung, Messzeitpunkt und Migrationshintergrund .....	236
Tabelle 43: Kennwerte der Interaktion zwischen Bedingung und Messzeitpunkt unter Kontrolle von Berufserfahrung als Kovariate .....	236
Tabelle 44: Kennwerte der Interaktion zwischen Bedingung und Messzeitpunkt unter Kontrolle der Anzahl der Kinder als Kovariate .....	237
Tabelle 45: Ergebnisse des Lack-of-Fit-Tests – Prädiktoren ohne linearen Zusammenhang .....	238
Tabelle 46: Korrelationen zwischen den Variablen sprachlicher Anregung und potenziellen Störvariablen .....	239

Tabelle 47: Kennwerte der hierarchischen Regressionsanalysen zur Vorhersage der sprachlichen Anregung durch die Skalen der "Feedback Orientation Scale" .....	240
Tabelle 48: Kennwerte der hierarchischen Regressionsanalysen zur Vorhersage der sprachlichen Anregung durch die Skalen des „AGQ-R“ .....	242
Tabelle 49: Kennwerte der hierarchischen Regressionen zur Vorhersage der sprachlichen Anregung durch die Berufsbezogene Motivation (LMI-K).....	244
Tabelle 50: Kennwerte der hierarchischen Regressionen zur Vorhersage der sprachlichen Anregung durch das Reflexionsniveau .....	245
Tabelle 51: Kennwerte der hierarchischen Regressionen zur Vorhersage der sprachlichen Anregung durch die Skalen des "Questionnaire for Training Evaluation" .....	246
Tabelle 52: Kennwerte der hierarchischen Regressionen zur Vorhersage der sprachlichen Anregung durch technische Probleme während des Feedbacks.....	248

## 12 Abbildungsverzeichnis

Abbildung 1: Faktoren und Dimensionen des "Classroom Assessment Scoring System" (CLASS Pre-K; Pianta, La Paro et al., 2008) .....	20
Abbildung 2: Kompetenzmodell nach Fröhlich-Gildhoff et al. (2011) .....	31
Abbildung 3: Indikatoren der Dimension "Language Modeling" der CLASS für ein sprachanregendes Verhalten .....	78
Abbildung 4: Konzept für ein online-basiertes Videofeedback .....	84
Abbildung 5: Ansicht vom "virtual team room" beim Feedbackgespräch .....	102
Abbildung 6: Baustein 1: Erfolgskriterien kommunizieren („Where am I going?“) .....	103
Abbildung 7: Baustein 2 ("How am I going?") und Baustein 3 ("Where to next?") .....	103
Abbildung 8: Baustein 4: Zielvereinbarung .....	104
Abbildung 9: Zielvereinbarung an Fachkraft schicken .....	105
Abbildung 10: Häufigkeit der verschiedenen technischen Probleme .....	133
Abbildung 11: Profildigramm Haupteffekt Messwiederholung .....	139
Abbildung 12: Profildigramm Haupteffekt Bedingung .....	140
Abbildung 13: Profildigramm Haupteffekt Bedingung – Geschlossene Fragen mit AM .....	147
Abbildung 14: Profildigramm Haupteffekt Messzeitpunkt – Geschlossene Fragen mit AM .....	147
Abbildung 15: Profildigramm Haupteffekt Bedingung – Ja/Nein-Fragen mit AM .....	148
Abbildung 16: Profildigramm Haupteffekt Messzeitpunkt – Ja/Nein-Fragen mit AM .....	148
Abbildung 17: Profildigramm Haupteffekt Bedingung – Offene Fragen mit AM .....	149
Abbildung 18: Profildigramm Haupteffekt Messzeitpunkt – Offene Fragen mit AM .....	149
Abbildung 19: Profildigramm Haupteffekt Bedingung – Offene Fragen ohne AM .....	150
Abbildung 20: Profildigramm Haupteffekt Messzeitpunkt – Offene Fragen ohne AM .....	150
Abbildung 21: Profildigramm Haupteffekt Bedingung – Erzählaufforderungen mit AM .....	151
Abbildung 22: Profildigramm Haupteffekt Messzeitpunkt – Erzählaufforderungen mit AM .....	151
Abbildung 23: Profildigramm Haupteffekt Bedingung – Erzählaufforderungen ohne AM .....	152
Abbildung 24: Profildigramm Haupteffekt Messzeitpunkt – Erzählaufforderungen ohne AM .....	152
Abbildung 25: Profildigramm Haupteffekt Bedingung – Anknüpfen .....	154
Abbildung 26: Profildigramm Haupteffekt Messzeitpunkt – Anknüpfen .....	154
Abbildung 27: Profildigramm Haupteffekt Bedingung – Wiederholung mit Korrektiv .....	158
Abbildung 28: Profildigramm Haupteffekt Messzeitpunkt – Wiederholung mit Korrektiv .....	158
Abbildung 29: Profildigramm Haupteffekt Bedingung – Wiederholung ohne Korrektiv .....	159
Abbildung 30: Profildigramm Haupteffekt Messzeitpunkt – Wiederholung ohne Korrektiv .....	159
Abbildung 31: Profildigramm Haupteffekt Bedingung – Erweiterung mit Korrektiv .....	159



Abbildung 32: Profildiagramm Haupteffekt Messzeitpunkt – Erweiterung mit Korrektiv .....	159
Abbildung 33: Profildiagramm Haupteffekt Bedingung – Erweiterung ohne Korrektiv .....	160
Abbildung 34: Profildiagramm Haupteffekt Messzeitpunkt – Erweiterung ohne Korrektiv .....	160

## 13 Abkürzungsverzeichnis

AGQ (-R)	Achievement Goal Questionnaire (Revised) (Instrument)
AQS	Attachment Q-Sort (Instrument)
BMFSFJ	Bundesministerium für Familie, Senioren, Frauen und Jugend
CIS	Caregiver Interaction Scale (Instrument)
CLASS	Classroom Assessment Scoring System (Instrument)
DO-RESI	Dortmunder Einschätzungsskala zur Erfassung sprachförderrelevanter Interaktionen (Instrument)
ECERS (-R)	Early Childhood Environment Rating Scale (Instrument)
EPPE	Effective Provision of Preschool Education (Projekt)
FOS	Feedback Orientation Scale (Instrument)
GiNA	Gestaltung von Interaktionsgelegenheiten im Alltag (Instrument)
HIT	Heidelberger Interaktionstraining
ICC	Intra-class correlation
IFP	Staatsinstitut für Frühpädagogik
KES (-R)	Kindergarten-Einschätzungsskala (deutsche Version der ECERS-R)
LMI-K	Leistungsmotivationsinventar (Kurzversion) (Instrument)
NUBBEK	Nationale Untersuchung zur Bildung, Betreuung und Erziehung in der frühen Kindheit (Projekt)
Q4TE	Questionnaire for Professional Training Evaluation (Instrument)
RC	Reliable Change
STRS	Student-Teacher-Relationship Scale (Instrument)
TILRS	Teacher Interaction and Language Rating Scale
VIG	Video Interaction Guidance
VSR	Videostimulierte Reflexion
WiFF	Weiterbildungsinitiative Frühpädagogische Fachkräfte (Projekt)

## 14 Anhang

Anhang A: Leitfaden für die Durchführung der Erhebungen

Anhang B: Kodier-Schema zur Einschätzung des sprachlichen Interaktionsverhaltens

Anhang C: Ergebnisse

- C1 Zielvereinbarungen
- C2 Anzahl an gestellten Fragen zu T1
- C3 Effekte auf die Anzahl sprachlicher Verhaltensweisen: Haupteffekt der Bedingung
- C4 Effekte auf die Anzahl sprachlicher Verhaltensweisen: Haupteffekt des Messzeitpunkts
- C5 Kontrolle von Störvariablen: Unterschiede zwischen Kontroll- und Feedbackgruppe
- C6 Kontrolle von Störvariablen: Einfluss auf die Effekte des Feedbacks
- C7 Regressionen: Lack-of-Fit-Test (Prädiktoren ohne linearen Zusammenhang)
- C8 Regressionen: Zusammenhang zwischen Kriterien und Störvariablen
- C9 Regressionen: Ergebnisse zur Vorhersage der Feedbackwirkung durch verschiedene Einflussfaktoren

Anhang A: Leitfaden für die Durchführung der Erhebungen

**Erhebungsprotokoll/-leitfaden**

	Erhebungsdatum: <span style="float: right;">Erheberin:</span>	
Name, Adresse und Telefonnummer der Einrichtung:		
Leitung	Name:	Code:
Name und Code der Fachkraft/Fachkräfte:	Name	Code
Anmerkungen		

## Vorbesprechung

### 1. Begrüßung und Besprechen des Ablaufs (Fachkräfte und evtl. Leitung)

- a. Kurzer Überblick über Ablauf
  - *Ich möchte gerne eine Videoaufnahme während einer Bilderbuchbetrachtung machen (Es wäre wichtig, dass bei unserem nächsten Besuch die Bilderbuchbetrachtung unter ähnlichen Bedingungen stattfindet (möglichst die gleichen Kinder und ähnliches Buch) -> sollen sich Namen der Kinder für t2 notieren!!!*
  - *Davor oder danach bräuchten wir noch etwa 30 Minuten Zeit, um das weitere Vorgehen zu besprechen und den schriftlichen Fragebogen auszufüllen.*
- b. Einwilligungserklärungen einsammeln (Eltern + Fachkräfte)
- c. Fragen?

### 2. Technik

- a. Technik-Check (Mikro-Check, Batterien-Check): Immer vorher prüfen, ob der Empfänger auch Ton empfängt (keine MUTE-Anzeige im Display oder über Kopfhörer prüfen) vor der Aufnahme sollten mindestens zwei Balken der Batterieanzeige (v.a. bei Sender prüfen, da ihr den später ggf. nicht mehr seht!) voll sein, ansonsten Batterien wechseln.
- b. Mikro-Übergabe an Fachkraft (ggf. später- je nachdem wie es Fachkraft lieber ist), "Mute" erklären (die Erzieherin kann das Mikro stumm schalten, muss sie aber nicht, da nach der Beobachtungszeit kein Ton aufgezeichnet wird)

### 3. Strukturfragebogen an Leitung geben (mit der Bitte, diesen bis zur zweiten Hospitation auszufüllen)

### 4. Ggf. Vorstellen bei den Kindern

## Hospitation

### 5. Anfertigen einer Videoaufnahme während Bilderbuchbetrachtung

#### Hinweise zur Durchführung der Video-Aufnahme

a) **Einwilligungen überprüfen:** Nur von Kindern, von denen die Einwilligung der Eltern vorliegt, dürfen Videoaufnahmen gemacht werden! Da die Videoaufnahme während dem Vorlesen gemacht werden soll, kann für diese Aktivität einfach die Gruppe an Kindern ausgewählt werden, von denen Einwilligungen vorliegen.

#### b) Mikro-Check

c) Auch mehrere Bilderbücher möglich, aber Achtung: nach BiBu-Betrachtung sind evtl. andere Kinder ohne Einwilligung im Bild, daher nach BiBu-Betrachtung Videoaufnahme beenden) während der Vorlese-/Bilderbuch-Situation durchlaufen lassen (möglichst keine Übergangszeit, sondern reine Vorlese-/Bilderbuch-Situation)

- Position gut wählen, nicht zu kleines Bild, aber **alle Beteiligten (Erzieherin + Kinder) müssen drauf sein:** Adressaten der Interaktionen mitfilmen (z.B. Vorlesen → Reaktionen der Kinder), in jedem Fall bei der Erzieherin und bei möglichst vielen Kindern **Mimik mit erfassen**
- Sich **möglichst wenig mit Handkamera bewegen** (Bewegung zieht Aufmerksamkeit auf Kamera), möglichst geringe Anpassungen der Aufnahmeposition eher selten vornehmen und Schwerpunkt auf Zoomen legen
- Aufnahmen **möglichst von den Rändern des Raumes aus vornehmen**, sich nicht unter die Kinder in die Mitte des Raumes bewegen (nur bei Wechsel von einer auf eine andere Seite des Raumes)
- **Distanz zur Interaktionssituation wahren** (je nach Raumgröße z.B. ~5-10 Meter)

#### Gebrauchsanweisung zur Technik

##### *Mikrofon und Taschensender*

- Mikrofon über der Brust mittig befestigen lassen (Vorsicht bei Schal, Kette, Haaren etc.)
- Kabel an Taschensender befestigen
- Taschensender einschalten:
  - links und rechts die Knöpfe leicht eindrücken → Deckel öffnet sich
  - on/off Taste länger gedrückt halten
  - Display: leuchtet orange und on-Knopf vorne leuchtet rot (wenn mute: keine Verbindung → mute Taste auf Taschensender verschieben oder Diversity-Empfänger einschalten)

- ggf. Batterien wechseln (immer morgens zwei Balken sicherstellen!): Klappe öffnen und Batterien austauschen

#### *Handkamera und Diversity-Empfänger*

- Stativ aufbauen (Hebel lösen und das Stativ auf die gewünschte Position einstellen, Hebel festdrücken)
- Kamera auf das Stativ aufsetzen (Hebel anziehen, Kamera von vorne nach hinten einführen und Hebel loslassen, Hebel fest andrücken)
- Diversity-Empfänger auf der Kamera befestigen: auf der Kamera in die Schienen einführen (so dass Antenne nach vorne absteht), an Drehrad festdrehen und Kabel rechts seitlich einstecken (Vorsicht: Das Kabel darf nicht vor der Linse hängen!)
- Diversity-Empfänger einschalten:
  - links und rechts die Knöpfe leicht eindrücken → Deckel öffnet sich
  - on/off Taste länger gedrückt halten
  - Display: leuchtet orange und on-Knopf vorne leuchtet rot (wenn mute: keine Verbindung → mute Taste auf Taschensender verschieben oder Taschensender einschalten)
- Kamera aufklappen
- Kamera einschalten: auf der linken Kameraseite on/off Taste drücken
- Filmkameramodus (seitlich rechts oben ist eine Filmkamera eingestellt, wenn nicht, diese einstellen)
- Kamera positionieren mit den 3 Hebeln am Stativ (zwei Hebel zum Drehen, ein Hebel zum Schwenken)
- Aufnahme starten: rote rec-Taste vorne an Kamera drücken, roter Punkt auf Display leuchtet
- Diversity-Empfänger auf Display der Kamera überprüfen: unten eine Leiste mit Ausschlagmesser (wenn er sich bewegt, ist Mikrophon an)
- ggf. Kamera von Stativ lösen und per Hand filmen (Hebel auf Stativ lösen und Kamera herabnehmen)
- Aufnahme stoppen: rote rec-Taste vorne an Kamera drücken
- Wiedergabemodus (seitlich rechts oben den Pfeil einstellen)
  - Per Touchscreen auf die Szene tippen, die man ansehen will (weitere Bedienung mit Stopp und Play)
  - Schließen des Wiedergabemodus (seitlich rechts die Filmkamera einstellen)
- Kamera ausschalten: links seitlich an Kamera on/off Taste drücken, Kamera schließen, Diversity-Empfänger ausschalten (Klappe öffnen, on/off Taste länger drücken) und abbauen (Kabel von Kamera lösen, Drehrad aufdrehen, aus der Schiene herausführen)

### **Nachbefragung** (nach Mittagessen)

Wie ist es Ihnen während der Beobachtung ergangen?  
Gab es heute irgendwelche Besonderheiten? War etwas anders als sonst?

Gibt es sonst noch irgendetwas, das Sie zum Tag/zur Beobachtung/zum Thema sagen möchten? Haben Sie Fragen an uns?

Bitte füllen Sie nun diesen **Fragebogen** aus (ca. 15 Min.) und stecken ihn anschließend ins Kuvert.

Info für nächsten Besuch (falls nicht schon in Vorbesprechung geklärt):  
Es wäre wichtig, dass bei unserem nächsten Besuch die Bilderbuchbetrachtung unter ähnlichen Bedingungen stattfindet (möglichst die gleichen Kinder und ähnliches Buch)

Während Fachkraft den Fragebogen ausfüllt, kann man vitero auf ihrem PC installieren und testen.

☐ WLAN oder ☐ LAN

Wurde vitero auf dem Kita-PC installiert? ☐ Ja ☐ Nein

Gab es technische Probleme? ☐ Ja ☐ Nein

Wenn ja, welche:

Herzlichen Dank für's Mitmachen  
Eine ausführliche Rückmeldung über die Beobachtung erhalten Sie nach Abschluss der Studie.

#### **Mitnehmen:**

- verschlossenes Kuvert mit ausgefülltem Fragebogen (Fachkraft)
- Einwilligungserklärungen



## Anhang B: Kodier-Schema zur Einschätzung des sprachlichen Interaktionsverhaltens

Kode	Definition/ Beschreibung	Beispiel
Längere Gespräche	Mind. 4-5 Sprecherwechsel mit einem Kind, nur wenn FK aktiv beteiligt ist; Äußerungen müssen Bezug aufeinander nehmen, sich auf ein Thema beziehen und eine inhaltliche Aussage enthalten (keine inhaltsfreien Hörsignale wie Hmm, Ok etc.).	
Geschlossene Fragen mit Antwortmöglichkeit	Arten von geschlossenen Fragen: - Erzieher kennt die Antwort und es gibt nur eine oder eine begrenzte Anzahl an richtigen/akzeptierten Antworten  - die Antwort ist nicht bekannt/offen, aber die Frage erfordert nur eine kurze Antwort - Entscheidungsfragen Codiert werden alle geschlossenen Fragen, bei denen die FK dem Kind die Möglichkeit/Zeit zu antworten gibt. V.a. Abwarten, keine sofortigen Nachfragen.	Welches Tier ist das?, Was ist das? Wo ist der Hase? Bei wem sind die Kinder? Was ist deine Lieblingsfarbe? Magst du lieber Hunde oder Katzen?
Geschlossene Fragen ohne Antwortmöglichkeit	Codiert werden alle geschlossenen Fragen, bei denen die FK dem Kind keine Möglichkeit/Zeit zu antworten gibt. Dies ist zum Beispiel der Fall, wenn Fachkraft direkt im Anschluss an die Frage selbst die Antwort sagt oder direkt danach eine andere Frage stellt. Zählt auch, wenn direkt im Anschluss eine Erzählaufforderung kommt.	<u>Was ist das?</u> Ein...
Ja/Nein-Fragen mit Antwortmöglichkeit	Fragen, auf die das Kind nur mit Ja oder Nein antworten kann. - Ja/Nein- Fragen - „Bestätigende“ Fragen: Fk macht eine Äußerung, auf die sie eine bestätigende Antwort erwartet. - Nachhaken bei falschen Antworten	Magst du Hunde?, Ist der Junge traurig? Das ist ein großer Baum, oder? Bist du sicher, dass...?
Ja/Nein-Fragen ohne Antwortmöglichkeit	Codiert werden alle Ja/Nein Fragen, bei denen die FK dem Kind keine Möglichkeit/Zeit zu antworten gibt. Dies ist zum Beispiel der Fall, wenn Fachkraft direkt im Anschluss an die Frage selbst die Antwort sagt oder direkt danach eine andere Frage stellt.	
Vermeintliche Fragen	Aussagen oder Aufforderungen, die als Frage formuliert, aber nicht als Frage gemeint sind- das heißt es wird nicht wirklich eine Antwort vom Kind erwartet. - Indirekte Verhaltensaufforderungen - Fragen, mit denen Kinder an Regeln erinnert werden. - Rhetorische Fragen	Kannst du dich bitte dort hinsetzen? Siehst du wie schön Max sitzt? Hab ich es euch nicht gesagt?

## Anhang

Offenen Fragen mit Antwortmöglichkeit	<ul style="list-style-type: none"> <li>- Fragen regen Antworten von mehr als einem Wort an (nicht Zwei-Wort-Äußerungen mit Artikel oder Präposition)</li> <li>- es gibt mehr als eine richtige/ akzeptierte Antwort</li> <li>- richtige Antwort ist evtl. bekannt, aber das Kind kann frei antworten</li> </ul> <p>Codiert werden alle offenen Fragen, bei denen die FK dem Kind die Möglichkeit/Zeit zu antworten gibt. V.a. Abwarten, keine sofortigen Nachfragen. Werden die Kinder zum Sprechen angeregt/aktiviert?</p> <p>Bei kombinierten Äußerungen zählt es als offene Frage, wenn am Ende der Äußerung ein Fragezeichen stehen würde.</p>	<p>Warum hat der Hase Angst vorm Fuchs?  Woran kannst du erkennen, dass...?  Wie könnten die Kinder das machen?  Was seht ihr auf dem Bild?  Fallen dir noch andere Tiere ein?  Was denkst du, was als Nächstes passiert?  Wie geht es dem Mädchen dann?  Was passiert auf dieser Seite?  Kannst du mir erzählen, was du auf der Seite siehst?  Erzähl du Daniel- was isst du gerne?</p>
Offenen Fragen ohne Antwortmöglichkeit	Codiert werden alle offenen Fragen, bei denen die FK dem Kind keine Möglichkeit/Zeit zu antworten gibt. Dies ist zum Beispiel der Fall, wenn Fachkraft direkt im Anschluss an die Frage selbst die Antwort sagt oder direkt danach eine andere Frage stellt. Zählt auch, wenn direkt im Anschluss eine Erzählaufforderung kommt.	
Aufforderungen zum Sprechen mit Antwortmöglichkeit	Codiert werden Aufforderungen/ Ermutigungen zum Sprechen, die keine Frageform aufweisen und bei denen das Kind Möglichkeit zum Antworten erhält. Können auch indirekte Sprech-Aufforderungen sein (z.B. Satz vervollständigen lassen), wenn deutlich ist, dass FK eine Antwort der Kinder erwartet (z.B. durch Sprechpause, Blickkontakt). Bei kombinierten Äußerungen zählt es als Erzählaufforderung, wenn am Ende der Äußerung ein Punkt oder Ausrufezeichen stehen würde.	<p>Erzählt mir mal, was ihr auf dem Bild seht.  Was ist das? <u>Ein...</u>  Was isst du gerne- erzähl doch mal!</p>
Aufforderungen zum Sprechen ohne Antwortmöglichkeit	Codiert werden Aufforderungen/ Ermutigungen zum Sprechen, die keine Frageform aufweisen und bei denen das Kind keine Möglichkeit zum Antworten erhält.	
Selftalk	Versprachlichen eigener Handlungen (wenn die Fachkraft parallel die Handlung ausführt).	<p>Ich blättere mal die Seite um.  Schauen wir mal, was auf der nächsten Seite passiert (wenn FK umblättert).</p>
Paralleltalk	Versprachlichen kindlicher Handlungen (wenn Kind/er die Handlung ausführen oder wenn es von allen gemeinsam gemacht wird); NICHT: Aufforderungen zum Handeln (z.B. „Tom setzt sich da hin“), in die Zukunft gerichtete Aussagen oder Aufforderungen zum Handeln (z.B. „Jetzt zählen wir mal die Affen auf dem Bild.“)	<p>Jetzt blättern wir die Seite um (wenn parallel gemeinsam umgeblättert wird).  Oh, du zeigst uns alle Schmetterlinge auf dem Bild.</p>
Wiederholung der kindlichen Äußerung ohne Korrektiv	Reine Wiederholung der kindlichen Äußerung, ohne Umformung oder Korrektur (es lag kein „Fehler“ in der kindlichen Äußerung vor)	K: 2 Katzen, Fk: (Ja-) 2 Katzen.
Korrektive Wiederholung der kindlichen Äußerung	Wiederholung der kindlichen Äußerung, wobei „Fehler“ reaktiv aufgenommen werden. Korrekturen können auf semantischer und/ oder syntaktisch-morphologischer Ebene erfolgen, z.B. inhaltlich falsche Antwort, falsch gebeugtes Verb oder falscher/fehlender Artikel.	<p>K: Die Baum, FK: Der Baum.  K: Baum, FK: Ein Baum.</p>

Korrektive Erweiterung	„Fehlerhafte“ kindliche Äußerungen werden aufgenommen und erweitert. Erweiterungen, Modellierungen oder Korrekturen können auf semantischer (zusätzl. Info) und/ oder syntaktisch-morphologischer Ebene erfolgen. Standardsprachliche Wiederholungen von dialektalen Äußerungen der Kinder werden nicht als Korrektiv gewertet.	K: Die Baum, FK: Der große Baum.
Erweiterung ohne Korrektiv	Äußerungen werden wiederholt und erweitert. Erweiterungen bzw. Modellierungen können auf semantischer (zusätzl. Info) und/ oder syntaktisch-morphologischer Ebene erfolgen; keine reinen Bestätigungen, wie z.B. „Das ist richtig“. Muss nicht innerhalb eines Satzes sein.	K: Ein Baum, FK: Ja genau- das ist ein Baum. K: Ein Baum. FK: Ein Baum. Der hat ganz viele Blätter.
Anknüpfen	Fachkraft reagiert auf eine Äußerung des Kindes <b>(wenn keiner der oben genannten Codes zutrifft: keine Frage, keine Wiederholung, keine Erweiterung)</b> ; nur inhaltliche Äußerungen mit Zusatzinformationen (NICHT: Hhm, Ok, Vielleicht, Kann sein etc.)	K: Was ist das? FK: Das ist eine Zange. K: Ich mag Spaghetti. FK: Ich mag lieber Pizza.

## Allgemeine Hinweise:

- jede Äußerung kann nur mit einem Code bewertet werden (Ausnahme: Längere Gespräche)
- Fragen werden auch gezählt, wenn sie im Buch vorkommen und die Fachkraft die Kinder nonverbal zum Antworten auffordert (z.B. deutliches Abwarten und Blickkontakt)
- Fehlen von Korrekativer Wiederholung (bei fehlerhaften Äußerungen des Kindes) werden nicht bewertet
- Auch organisatorische Themen werden bei Fragen etc. mitgezählt

Achtung: Wenn bei einer Äußerung zwei Codes zutreffen (z.B. eine Wiederholung, die als Frage intoniert ist) wird immer der höherwertige Code vergeben! Beispiel: K sagt „Ich nicht“, Fk reagiert mit „Du nicht?“ -> zählt als Wiederholung, da dies höher gewertet wird als Ja/Nein-Fragen.

## Ranking:

1. Offene Fragen	5. Erweiterung ohne Korrektiv
2. Erzählaufforderungen	6. a) Wiederholung ohne Korrektiv
3. Geschlossene Fragen	b) Ja/Nein-Fragen
4. a) Korrektive Erweiterung b) Korrektive Wiederholung	7. Vermeintliche Fragen

Anhang C: ErgebnisseC1 Zielvereinbarungen

Tabelle 36: Zielvereinbarungen

<b>Zielformulierungen</b>	<b>Absolute Häufigkeit</b>
Offene Fragen stellen	38
Mehr Dialoge führen	16
Auf Äußerungen der Kinder eingehen	15
Weniger vorwegnehmen und häufiger Kinder sprechen lassen	14
Kinder mehr einbeziehen	12
Den Kindern mehr Zeit zum Sprechen lassen	12
Weniger vorlesen, mehr mit Kindern über Buch sprechen/Kinder erzählen lassen	12
Äußerungen wiederholen	9
Begriffe gemeinsam erläutern	6
Kindern die Möglichkeit für eigene Ideen lassen/Mehr auf Interessen der Kinder eingehen	4
Mehr Pausen zum erzählen machen	4
Bezug zu Alltag der Kinder herstellen	4
Kinder zum Zusammenfassen/Wiederholen anregen	2
Langsamer sprechen	2
Auf die eigene Sprache achten	2
reiches Materialangebot	2
Kindern Möglichkeit für Fragen lassen	1
Buchseiten abkleben	1
Dialog unter den Kindern fördern	1
Noch mehr auf stille Kinder eingehen	1
Fragen von Kindern an andere Kinder weitergeben	1

C2 Anzahl an gestellten Fragen zu T1

Tabelle 37: Häufigkeit der verschiedenen Frageformen zu T1

Frageformen	N	Summe	Häufigkeit in Prozent	
Geschlossene Fragen mit Antwortmöglichkeit (AM)_T1	96	1278	22.08%	26.42%
Geschlossene Fragen ohne Antwortmöglichkt (AM)_T1	96	251	4.34%	
Ja/Nein-Fragen mit AM_T1	96	2059	35.7%	41.66%
Ja/Nein-Fragen ohne AM_T1	96	352	6.08%	
Vemeintliche Fragen_T1	96	266	4.60%	
Offene Fragen mit AM_T1	96	1312	22.67%	27.33%
Offene Fragen ohne AM_T1	96	270	4.66%	
Anzahl an gestellten Fragen insgesamt	96	5788	100%	

C3 Effekte auf die Anzahl sprachlicher Verhaltensweisen: Haupteffekt der Bedingung

Tabelle 38: Kennwerte für den Haupteffekt der Bedingung in Bezug auf das sprachliche Verhalten

	<i>F</i>	<i>p</i>	$\eta_p^2$
Längere Gespräche	3.97	0.05	0.04
Geschlossene Fragen mit AM	1.72	0.19	0.02
Geschlossene Fragen ohne AM	0.11	0.74	0.00
Ja/Nein-Fragen mit AM	3.52	0.06	0.04
Ja/Nein-Fragen ohne AM	0.06	0.80	0.00
Vermeintliche Fragen	0.01	0.92	0.00
Offene Fragen mit AM	4.74	0.03	0.05
Offene Fragen ohne AM	3.42	0.07	0.04
Erzählaufforderungen mit AM	0.59	0.44	0.01
Erzählaufforderungen ohne AM	0.02	0.90	0.00
Selftalk	0.72	0.40	0.01
Paralleltalk	0.00	0.95	0.00
Wiederholung mit Korrektur	5.74	0.02	0.06
Wiederholung ohne Korrektur	4.94	0.03	0.05
Erweiterung mit Korrektur	3.22	0.08	0.03
Erweiterung ohne Korrektur	4.00	0.05	0.04
Anknüpfen	0.87	0.35	0.01

Anmerkung:  $df(\text{Bedingung}) = 1$ ,  $df(\text{Fehler}) = 94$

C4 Effekte auf die Anzahl sprachlicher Verhaltensweisen: Haupteffekt des Messzeitpunkts

Tabelle 39: Kennwerte für den Haupteffekt des Messzeitpunkts in Bezug auf das sprachliche Verhalten

	<i>F</i>	<i>p</i>	$\eta_p^2$
Längere Gespräche	6.55	0.01	0.07
Geschlossene Fragen mit AM	7.34	0.01	0.07
Geschlossene Fragen ohne AM	1.87	0.18	0.02
Ja/Nein-Fragen mit AM	2.95	0.09	0.03
Ja/Nein-Fragen ohne AM	0.55	0.46	0.01
Vermeintliche Fragen	0.01	0.93	0.00
Offene Fragen mit AM	17.85	0.00	0.16
Offene Fragen ohne AM	7.69	0.01	0.08
Erzählaufforderungen mit AM	6.21	0.01	0.06
Erzählaufforderungen ohne AM	0.03	0.86	0.00
Selftalk	0.78	0.38	0.01
Paralleltalk	0.07	0.79	0.00
Wiederholung mit Korrektur	3.64	0.06	0.04
Wiederholung ohne Korrektur	7.15	0.01	0.07
Erweiterung mit Korrektur	8.72	0.00	0.09
Erweiterung ohne Korrektur	16.19	0.00	0.15
Anknüpfen	5.46	0.02	0.06

Anmerkung:  $df(\text{Bedingung}) = 1$ ,  $df(\text{Fehler}) = 94$

C5 Kontrolle von Störvariablen: Unterschiede zwischen Kontroll- und Feedbackgruppe

Tabelle 40: Kennwerte der Chi-Quadrat-Tests zur Ermittlung von Unterschieden in kategorialen Störvariablen

	$\chi^2$	$df$	$p$	$\phi$
Geschlecht	0.21	1	1.00	0.05
Migrationshintergrund	5.04	1	0.04	0.23
Mehrsprachigkeit	1.79	1	0.25	0.14
Höchster Abschluss	4.94	4	0.28	0.23
Höchste Position	1.13	4	0.90	0.11



Tabelle 41: Kennwerte der t-Tests für unabhängige Stichproben zur Ermittlung von Unterschieden in metrischen Störvariablen

	<i>t</i>	<i>df</i>	<i>p</i>	<i>g</i>	
Alter der Fachkraft	0.84	90	0.40	0.18	
Berufserfahrung in der Kita	0.19	90	0.85	0.04	
Berufserfahrung als Fachkraft	1.87	90	0.06	0.39	
Leistung T1	Language Modeling	1.09	94	0.28	0.22
	Geschlossene Fragen mit AM	0.37	94	0.71	0.08
	Ja/Nein- Fragen mit AM	-0.23	94	0.82	0.05
	Offene Fragen mit AM	0.41	94	0.68	0.08
	Offene Fragen ohne AM	0.11	94	0.91	0.02
	Erzählaufforderungen mit AM	1.11	94	0.27	0.23
	Wiederholung mit Korrektur	-0.24	94	0.81	0.05
	Wiederholung ohne Korrektur	-0.15	94	0.88	0.03
	Erweiterung mit Korrektur <sup>1</sup>	-0.69	63.57	0.49	0.15
	Erweiterung ohne Korrektur	0.54	94	0.59	0.11
	Anknüpfen	0.18	94	0.85	0.04
	Berufsbezogene Motivation	-0.72	90	0.47	0.15
	Unterschiede zwischen T1 und T2	Anzahl der Kinder	1.87	81	0.07
Altersspanne der Kinder		0.96	80	0.34	0.21
Anzahl Kinder nicht Deutsch sprechend <sup>1</sup>		-0.76	55.94	0.45	0.17
Anzahl Kinder nicht Deutsch sprechend mit DAZ		-0.47	80	0.64	0.10
Anzahl Kinder Sprachanfänger		0.99	79	0.32	0.22
Anzahl Kinder Sprachanfänger mit DAZ <sup>1</sup>		-0.16	67	0.88	0.03
Anzahl Kinder spracherfahren <sup>1</sup>		-0.01	59	0.99	0.00
Anzahl Kinder redegewandt		1.28	79	0.20	0.29
Anzahl Kinder Migration mit Sprachauffälligkeiten		0.39	79	0.70	0.09
Abstand zwischen T1 und T2 <sup>1</sup>		-0.89	55.25	0.38	0.20

Anmerkung: <sup>1</sup> Aufgrund inhomogener Varianzen wurde auf die Kennwerte des Welch-Tests mit entsprechender Korrektur der Freiheitsgrade zurückgegriffen.

C6 Kontrolle von Störvariablen: Einfluss auf die Effekte des Feedbacks

Migrationshintergrund der Fachkräfte:

Tabelle 42: Kennwerte der Interaktion zwischen Bedingung, Messzeitpunkt und Migrationshintergrund

	<i>F</i>	<i>p</i>	$\eta_p^2$
Language Modeling	0.01	0.94	0.00
Geschlossene Fragen mit AM	0.92	0.34	0.01
Ja/Nein-Fragen mit AM	0.13	0.72	0.00
Offene Fragen mit AM	0.01	0.91	0.00
Offene Fragen ohne AM	0.41	0.52	0.00
Erzählaufforderungen mit AM	0.26	0.61	0.00
Wiederholung mit Korrektur	0.07	0.80	0.00
Wiederholung ohne Korrektur	0.23	0.63	0.00
Erweiterung mit Korrektur	0.02	0.89	0.00
Erweiterung ohne Korrektur	0.64	0.43	0.01
Anknüpfen	1.17	0.28	0.01

Anmerkung: df (Interaktion) = 1, df (Fehler) = 88

Berufserfahrung als Fachkraft:

Tabelle 43: Kennwerte der Interaktion zwischen Bedingung und Messzeitpunkt unter Kontrolle von Berufserfahrung als Kovariate

	<i>F</i>	<i>p</i>	$\eta_p^2$
Language Modeling	24.05	< 0.01	0.21
Geschlossene Fragen mit AM	5.72	0.02	0.06
Ja/Nein-Fragen mit AM	10.80	< 0.01	0.11
Offene Fragen mit AM	37.64	< 0.01	0.30
Offene Fragen ohne AM	11.83	< 0.01	0.12
Erzählaufforderungen mit AM	6.37	0.01	0.07
Wiederholung mit Korrektur	7.93	0.01	0.08
Wiederholung ohne Korrektur	19.75	< 0.01	0.18
Erweiterung mit Korrektur	4.46	0.04	0.05
Erweiterung ohne Korrektur	24.52	< 0.01	0.22
Anknüpfen	7.44	0.01	0.08

Anmerkung: df (Interaktion) = 1, df (Fehler) = 89

## Differenz in der Anzahl der Kinder bei der Bilderbuchbetrachtung:

Tabelle 44: Kennwerte der Interaktion zwischen Bedingung und Messzeitpunkt unter Kontrolle der Anzahl der Kinder als Kovariate

	<i>F</i>	<i>p</i>	$\eta_p^2$
Language Modeling	16.70	< 0.01	0.17
Geschlossene Fragen mit AM	4.31	0.04	0.05
Ja/Nein-Fragen mit AM	6.54	0.01	0.08
Offene Fragen mit AM	22.97	< 0.01	0.22
Offene Fragen ohne AM	9.66	< 0.01	0.11
Erzählaufforderungen mit AM	7.75	0.01	0.09
Wiederholung mit Korrektur	4.34	0.04	0.05
Wiederholung ohne Korrektur	14.12	< 0.01	0.15
Erweiterung mit Korrektur	4.03	0.05	0.05
Erweiterung ohne Korrektur	20.36	< 0.01	0.20
Anknüpfen	5.33	0.02	0.06

Anmerkung: df (Interaktion) = 1, df (Fehler) = 80

C7 Regressionen: Lack-of-Fit-Test (Prädiktoren ohne linearen Zusammenhang)

In der folgenden Tabelle sind die jeweiligen Prädiktoren aufgelistet, bei denen kein linearer Zusammenhang zum Kriterium festgestellt wurde und welche daraufhin aus den Regressionsanalysen ausgeschlossen wurden.

Tabelle 45: Ergebnisse des Lack-of-Fit-Tests – Prädiktoren ohne linearen Zusammenhang

	<b>Geschlossene Fragen mit AM</b>	<b>Ja/Nein Fragen mit AM</b>	<b>Offene Fragen mit AM</b>	<b>Erzählauforderungen mit AM</b>	<b>Wiederholung mit Korrektur</b>	<b>Wiederholung ohne Korrektur</b>	<b>Anknüpfen</b>
Prädiktoren, zu denen kein linearer Zusammenhang besteht	Performance Avoidance	Leistung T1	Position der Fachkraft	Mastery approach	Satisfaction	Performance Avoidance	Mastery avoidance
		Performance Avoidance		Performance approach	Utility		
					Knowledge		
					LMI-K		
					Mastery approach		
					Mastery avoidance		
					FOS Utility		
					FOS Accountability		
					FOS Social Awareness		

## C8 Regressionen: Zusammenhang zwischen Kriterien und Störvariablen

Tabelle 46: Korrelationen zwischen den Variablen sprachlicher Anregung und potenziellen Störvariablen

		Diff	T1	Alter	Jahre in Einrichtung	Feedback- geber	Jahre als Fachkraft	Mehr- sprachigkeit	Fachschule Kinderpfl.	Fachschule Sozialpäd.	Fachhoch- schule	Universität	Ergänzungs- kraft	Fachkraft ol.	Fachkraft mL	stellv Leitung	Leitung	Migration Großeltern	Migration Eltern	Migration Fk	Zeit zwischen Feedback und T2
LM_gesamt_ Diff	Pearson- Korrelation	1	-.460**	,068	,090	,068	,258	,072	-.387**	,230	,114	,002	-.273	-.143	-.017	,251	,264	,069	,039	,028	-.044
	Sig. (2-seitig)		,001	,659	,562	,645	,091	,643	,010	,138	,467	,990	,088	,379	,917	,119	,100	,656	,802	,858	,765
	N	48	48	44	44	44	44	44	43	43	43	43	40	40	40	40	40	44	44	44	48
Geschl_m_D iff	Pearson- Korrelation	1	-.338**	,124	,090	,172	-.047	,017	-.196	-.089	,344*	-.010	-.140	-.057	,029	,262	-.167	,195	-.037	-.005	-.151
	Sig. (2-seitig)		,019	,422	,563	,242	,760	,915	,208	,570	,024	,950	,388	,729	,858	,103	,303	,205	,812	,976	,304
	N	48	48	44	44	48	44	44	43	43	43	43	40	40	40	40	40	44	44	44	48
JaNein_m_D iff	Pearson- Korrelation	1	-.670**	,020	,255	-.019	,230	-.063	-.136	,104	,080	-.089	-.199	,320*	-.295	,077	,207	-.156	,165	-.148	-.113
	Sig. (2-seitig)		,000	,898	,094	,899	,133	,683	,384	,507	,611	,569	,218	,044	,064	,635	,200	,312	,283	,337	,444
	N	48	48	44	44	48	44	44	43	43	43	43	40	40	40	40	40	44	44	44	48
Offen_m_Diff	Pearson- Korrelation	1	-.186	,090	,251	,169	,225	,051	-.152	,236	-.110	-.053	-.186	,030	-.288	,170	,469**	-.098	-.075	-.091	-.176
	Sig. (2-seitig)		,205	,559	,101	,251	,143	,742	,332	,127	,482	,736	,251	,855	,072	,295	,002	,527	,629	,559	,232
	N	48	48	44	44	48	44	44	43	43	43	43	40	40	40	40	40	44	44	44	48
Offen_o_Diff	Pearson- Korrelation	1	-.116	,182	,102	-.116	,245	,025	-.149	,072	-.118	,236	,128	,143	-.315*	,068	,029	-.163	-.020	,132	,083
	Sig. (2-seitig)		,433	,236	,511	,432	,109	,872	,339	,647	,450	,127	,430	,379	,047	,675	,860	,289	,895	,394	,577
	N	48	48	44	44	48	44	44	43	43	43	43	40	40	40	40	40	44	44	44	48
Erz_auff_m_ Diff	Pearson- Korrelation	1	,046	,295	,201	,112	,220	-.283	-.151	,165	-.116	,074	-.192	-.183	-.146	,453**	,105	-.134	-.065	-.154	-.107
	Sig. (2-seitig)		,756	,052	,191	,447	,151	,063	,335	,290	,460	,636	,236	,258	,368	,003	,518	,386	,676	,319	,470
	N	48	48	44	44	48	44	44	43	43	43	43	40	40	40	40	40	44	44	44	48
Wiedh_mK_ Diff	Pearson- Korrelation	1	-.285*	,091	,058	,257	,115	-.238	-.255	,158	-.035	,131	-.216	-.057	,175	,106	-.057	-.031	-.099	-.005	-.037
	Sig. (2-seitig)		,049	,555	,710	,078	,458	,119	,098	,313	,825	,401	,180	,726	,280	,516	,728	,841	,524	,976	,803
	N	48	48	44	44	48	44	44	43	43	43	43	40	40	40	40	40	44	44	44	48
Wiedh_oK_D iff	Pearson- Korrelation	1	-.336*	,151	,160	,297*	,143	-.143	-.311*	,235	,082	-.073	-.293	-.140	-.094	,385*	,210	,134	-.025	-.171	-.258
	Sig. (2-seitig)		,020	,328	,300	,040	,354	,353	,042	,129	,601	,643	,066	,389	,566	,014	,193	,385	,874	,268	,076
	N	48	48	44	44	48	44	44	43	43	43	43	40	40	40	40	40	44	44	44	48
Erw_mK_Diff	Pearson- Korrelation	1	,180	,207	,017	,104	-.064	-.158	-.180	,216	-.107	,018	-.108	-.177	,170	,036	,080	-.067	-.111	-.051	-.023
	Sig. (2-seitig)		,222	,177	,912	,480	,682	,306	,247	,165	,497	,911	,509	,275	,295	,827	,622	,666	,473	,744	,874
	N	48	48	44	44	48	44	44	43	43	43	43	40	40	40	40	40	44	44	44	48
Erw_oK_Diff	Pearson- Korrelation	1	-.141	,038	,057	,058	,187	-.233	-.219	-.012	,321*	-.080	-.216	-.013	-.103	,326*	,009	-.243	-.018	-.104	-.056
	Sig. (2-seitig)		,340	,809	,713	,698	,225	,127	,157	,940	,036	,610	,181	,934	,529	,040	,954	,112	,908	,502	,706
	N	48	48	44	44	48	44	44	43	43	43	43	40	40	40	40	40	44	44	44	48
Ankn_Diff	Pearson- Korrelation	1	,059	,039	,043	-.339*	,232	,116	-.051	,097	-.037	-.048	-.182	,084	-.185	,318*	-.034	-.126	,002	,191	,066
	Sig. (2-seitig)		,691	,803	,783	,018	,130	,454	,747	,537	,815	,758	,261	,607	,253	,045	,836	,414	,989	,215	,658
	N	48	48	44	44	48	44	44	43	43	43	43	40	40	40	40	40	44	44	44	48

## C9 Regressionen: Ergebnisse zur Vorhersage der Feedbackwirkung durch verschiedene Einflussfaktoren

Tabelle 47: Kennwerte der hierarchischen Regressionsanalysen zur Vorhersage der sprachlichen Anregung durch die Skalen der "Feedback Orientation Scale"

	Prädiktoren	Modell 1			Modell 2		
		B	SE B	p	B	SE B	p
Language Modeling <sup>1</sup>	Language Modeling T1	-0.40	0.11	<0.01	-0.45	0.11	<0.01
	Abschluss: Kinderpflege	-0.72	0.42	0.10	-0.89	0.45	0.06
	FOS_Utility				-0.08	0.31	0.74
	FOS_Accountability				-0.38	0.31	0.15
	FOS_Social Awareness				0.51	0.23	0.04
	FOS_Self Efficacy				-0.17	0.17	0.25
		$R^2 = 0.34$			$\Delta R^2 = 0.11, p = 0.17$		
Geschlossene Fragen mit AM	Geschlossene Fragen mit AM_T1	-0.44	0.19	0.03	-0.59	0.20	0.01
	Abschluss: Fachhochschule	11.89	5.61	0.04	14.57	5.91	0.02
	FOS_Utility				7.46	4.78	0.13
	FOS_Accountability				-4.95	4.94	0.32
	FOS_Social Awareness				4.36	3.72	0.25
	FOS_Self Efficacy				0.74	2.81	0.79
		$R^2 = 0.22$			$\Delta R^2 = 0.10, p = 0.28$		
Ja/Nein-Fragen mit AM	Ja/Nein-Fragen mit AM_T1	-0.68	0.10	<0.01	-0.68	0.11	<0.01
	Position: FK ohne Gruppenleitung	9.25	4.67	0.06	9.14	5.08	0.08
	FOS_Utility				0.37	4.43	0.93
	FOS_Accountability				-1.58	4.78	0.74
	FOS_Social Awareness				2.00	3.77	0.60
	FOS_Self Efficacy				1.30	2.63	0.63
		$R^2 = 0.58$			$\Delta R^2 = 0.01, p = 0.95$		
Offene Fragen mit AM	Position: Leitung	21.69	6.63	<0.01	21.14	6.68	<0.01
	FOS_Utility				3.64	4.00	0.37
	FOS_Accountability				-5.55	4.18	0.19
	FOS_Social Awareness				-0.05	3.33	0.99
	FOS_Self Efficacy				-2.77	2.31	0.24
		$R^2 = 0.22$			$\Delta R^2 = 0.09, p = 0.40$		
Offene Fragen ohne AM	Position: Gruppenleitung	-2.95	1.44	0.05	-2.74	1.46	0.07
	FOS_Utility				-3.79	1.44	0.01
	FOS_Accountability				2.19	1.52	0.16
	FOS_Social Awareness				1.43	1.26	0.26
	FOS_Self Efficacy				-0.53	0.85	0.53
		$R^2 = 0.1$			$\Delta R^2 = 0.16, p = 0.14$		
Erzählaufforderungen mit AM	Position: stellv. Leitung	15.86	5.06	<0.01	15.57	5.17	0.01
	FOS_Utility				-1.86	4.83	0.70

		Modell 1			Modell 2		
	Prädiktoren	B	SE B	p	B	SE B	p
	FOS_Accountability				-3.92	5.10	0.45
	FOS_Social Awareness				7.01	4.06	0.09
	FOS_Self Efficacy				-2.38	2.83	0.41
		$R^2 = 0.21$			$\Delta R^2 = 0.08, p = 0.45$		
Wiederholung mit Korrektur	Wiederholung mit Korrektur_T1	-0.68	0.33	0.05	-0.67	0.36	0.07
	FOS_Self Efficacy				-0.08	0.78	0.92
		$R^2 = 0.09$			$\Delta R^2 = 0.00, p = 0.92$		
Wiederholung ohne Korrektur	Wiederh. ohne Korrektur_T1	-0.43	0.17	0.02	-0.42	0.17	0.02
	Feedbackgeber	7.70	3.91	0.06	8.47	3.99	0.04
	Abschluss: Kinderpflege	-3.13	5.13	0.55	-8.10	5.82	0.17
	Position: stellv. Leitung	9.40	4.71	0.05	7.48	4.86	0.13
	FOS_Utility				6.07	4.18	0.16
	FOS_Accountability				-8.61	4.42	0.06
	FOS_Social Awareness				6.79	3.98	0.10
	FOS_Self Efficacy				-1.09	2.48	0.66
		$R^2 = 0.39$			$\Delta R^2 = 0.10, p = 0.21$		
Erweiterungen mit Korrektur	FOS_Utility	-1.28	1.28	0.32			
	FOS_Accountability	0.15	1.34	0.91			
	FOS_Social Awareness	0.77	0.96	0.43			
	FOS_Self Efficacy	-0.08	0.73	0.92			
		$R^2 = 0.04$					
Erweiterungen ohne Korrektur	Abschluss: Fachhochschule	9.71	3.67	0.01	11.40	4.40	0.01
	Position: stellv. Leitung	8.09	2.91	0.01	8.61	2.98	0.01
	FOS_Utility				0.27	2.73	0.92
	FOS_Accountability				-3.83	2.86	0.19
	FOS_Social Awareness				3.79	2.29	0.11
	FOS_Self Efficacy				1.08	1.81	0.56
		$R^2 = 0.25$			$\Delta R^2 = 0.08, p = 0.41$		
Anknüpfen	Feedbackgeber	-3.67	1.84	0.05	-3.03	2.01	0.14
	Position: stellv. Leitung	5.09	2.18	0.03	4.75	2.28	0.05
	FOS_Utility				-0.86	2.11	0.69
	FOS_Accountability				-0.60	2.23	0.79
	FOS_Social Awareness				2.76	1.82	0.14
	FOS_Self Efficacy				-0.44	1.24	0.73
		$R^2 = 0.19$			$\Delta R^2 = 0.05, p = 0.67$		

Anmerkung: <sup>1</sup> Aufgrund von Heteroskedastizität erfolgte eine korrigierte Signifikanzprüfung mithilfe des Bootstrap-Verfahrens.

Tabelle 48: Kennwerte der hierarchischen Regressionsanalysen zur Vorhersage der sprachlichen Anregung durch die Skalen des „AGQ-R“

	Prädiktoren	Modell 1			Modell 2		
		B	SE B	p	B	SE B	p
Language Modeling	Language Modeling T1	-0.40	0.12	<0.01	-0.40	0.13	<0.01
	Abschluss: Kinderpflege	-0.72	0.32	0.03	-0.66	0.37	0.08
	Mastery approach goals				-0.02	0.29	0.94
	Mastery avoidance goals				0.00	0.18	0.99
	Performance approach goals				0.11	0.18	0.55
	Performance avoidance goals				-0.08	0.16	0.62
		$R^2 = 0.34$			$\Delta R^2 = 0.01, p = 0.98$		
Geschlossene Fragen mit AM	Geschlossene Fragen mit AM_T1	-0.44	0.19	0.03	-0.6	0.2	0.01
	Abschluss: Fachhochschule	11.89	5.61	0.04	13.9	5.4	0.01
	Mastery approach goals				5.0	4.1	0.23
	Mastery avoidance goals				2.2	2.5	0.39
	Performance approach goals				1.0	2.1	0.65
		$R^2 = 0.22$			$\Delta R^2 = 0.13, p = 0.08$		
Ja/Nein-Fragen mit AM	Ja/Nein-Fragen mit AM_T1	-0.68	0.10	<0.01	-0.72	0.11	<0.01
	Position: FK ohne Gruppenleitung	9.25	4.67	0.06	8.68	4.86	0.08
	Mastery approach goals				1.37	3.98	0.73
	Mastery avoidance goals				-2.07	2.40	0.39
	Performance approach goals				3.63	2.16	0.10
		$R^2 = 0.58$			$\Delta R^2 = 0.04, p = 0.30$		
Offene Fragen mit AM	Position: Leitung	21.69	6.63	<0.01	21.23	7.09	0,01
	Mastery approach goals				0.34	3.88	0,93
	Mastery avoidance goals				0.87	2.37	0,72
	Performance approach goals				2.12	2.51	0,41
	Performance avoidance goals				-1.33	2.12	0,54
		$R^2 = 0.22$			$\Delta R^2 = 0.02, p = 0.90$		
Offene Fragen ohne AM	Position: Gruppenleitung	-2.95	1.44	0.05	-2.78	1.48	0.07
	Mastery approach goals				-2.30	1.40	0.11
	Mastery avoidance goals				1.39	0.86	0.11
	Performance approach goals				0.75	0.91	0.42
	Performance avoidance goals				-0.26	0.76	0.73
		$R^2 = 0.1$			$\Delta R^2 = 0.16, p = 0.14$		
Erzählaufforderungen mit AM	Position: stellv. Leitung	15.86	5.06	<0.01	16.87	5.27	<0.01
	Mastery avoidance goals				-1.28	2.60	0.63
	Performance avoidance goals				1.83	1.92	0.35
		$R^2 = 0.21$			$\Delta R^2 = 0.02, p = 0.64$		
Wiederholung mit Korrektur	Wiederholung mit Korrektur_T1	-0.68	0.33	0.05	-0.57	0.33	0.10
	Performance approach goals				-0.59	0.73	0.42
	Performance avoidance goals				-0.27	0.57	0.64
		$R^2 = 0.09$			$\Delta R^2 = 0.06, p = 0.28$		



		Modell 1			Modell 2		
	Prädiktoren	B	SE B	p	B	SE B	p
Wiederholung ohne Korrektur	Wiederh. ohne Korrektur_T1	-0.43	0.17	0.02	-0.45	0.18	0.02
	Feedbackgeber	7.70	3.91	0.06	9.35	4.57	0.05
	Abschluss: Kinderpflege	-3.13	5.13	0.55	-4.73	5.45	0.39
	Position: stellv. Leitung	9.40	4.71	0.05	8.02	4.97	0.12
	Mastery approach goals				1.91	4.05	0.64
	Mastery avoidance goals				0.51	2.52	0.84
	Performance approach goals				1.29	2.59	0.62
		$R^2 = 0.39$			$\Delta R^2 = 0.03, p = 0.72$		
Erweiterungen mit Korrektur	Mastery approach goals	0.25	1.18	0.83			
	Mastery avoidance goals	0.33	0.72	0.65			
	Performance approach goals	-0.04	0.74	0.95			
	Performance avoidance goals	-0.29	0.63	0.65			
		$R^2 = 0.01$					
Erweiterungen ohne Korrektur	Abschluss: Fachhochschule	9.71	3.67	0.01	10.29	4.10	0.02
	Position: stellv. Leitung	8.09	2.91	0.01	8.61	3.19	0.01
	Mastery approach goals				-0.52	2.74	0.85
	Mastery avoidance goals				0.88	1.64	0.59
	Performance approach goals				-1.40	1.74	0.43
	Performance avoidance goals				0.75	1.50	0.62
		$R^2 = 0.25$			$\Delta R^2 = 0.02, p = 0.91$		
Anknüpfen	Feedbackgeber	-3.67	1.84	0.05	-3.28	2.12	0.13
	Position: stellv. Leitung	5.09	2.18	0.03	4.47	2.29	0.06
	Mastery approach goals				2.93	1.80	0.11
	Performance approach goals				-0.06	1.38	0.97
	Performance avoidance goals				-0.31	1.06	0.78
		$R^2 = 0.19$			$\Delta R^2 = 0.08, p = 0.34$		

Tabelle 49: Kennwerte der hierarchischen Regressionen zur Vorhersage der sprachlichen Anregung durch die Berufsbezogene Motivation (LMI-K)

	Prädiktoren	Modell 1			Modell 2		
		B	SE B	p	B	SE B	p
Language Modeling	Language Modeling T1	-0.40	0.12	<0.01	-0.41	0.12	<0.01
	Abschluss: Kinderpflege	-0.72	0.32	0.03	-0.73	0.33	0.03
	Berufsbezogene Motivation				0.00	0.01	0.88
		$R^2 = 0.34$			$\Delta R^2 = 0.00, p = 0.88$		
Geschlossene Fragen mit AM	Geschlossene Fragen mit AM_T1	-0.44	0.19	0.03	-0.53	0.20	0.01
	Abschluss: Fachhochschule	11.89	5.61	0.04	13.03	5.62	0.03
	Berufsbezogene Motivation				0.14	0.11	0.19
		$R^2 = 0.22$			$\Delta R^2 = 0.03, p = 0.19$		
Ja/Nein-Fragen mit AM	Ja/Nein-Fragen mit AM_T1	-0.68	0.10	<0.01	-0.74	0.11	<0.01
	Position: FK ohne Gruppenleitung	9.25	4.67	0.06	9.81	4.48	0.04
	Berufsbezogene Motivation				0.19	0.09	0.05
		$R^2 = 0.58$			$\Delta R^2 = 0.05, p = 0.05$		
Offene Fragen mit AM	Position: Leitung	21.69	6.63	<0.01	21.69	6.72	<0.01
	Berufsbezogene Motivation				-0.01	0.09	0.91
		$R^2 = 0.22$			$\Delta R^2 = 0.00, p = 0.91$		
Offene Fragen ohne AM	Position: Gruppenleitung	-2.95	1.44	0.05	-3.11	1.48	0.04
	Berufsbezogene Motivation				-0.02	0.03	0.56
		$R^2 = 0.1$			$\Delta R^2 = 0.01, p = 0.56$		
Erzählaufforderungen mit AM	Position: stellv. Leitung	15.86	5.06	<0.01	16.89	5.22	<0.01
	Berufsbezogene Motivation				-0.09	0.11	0.40
		$R^2 = 0.21$			$\Delta R^2 = 0.02, p = 0.40$		
Wiederholung ohne Korrektur	Wiederh. ohne Korrektur_T1	-0.43	0.17	0.02	-0.44	0.18	0.02
	Feedbackgeber	7.70	3.91	0.06	7.88	4.12	0.06
	Abschluss: Kinderpflege	-3.13	5.13	0.55	-3.40	5.45	0.54
	Position: stellv. Leitung	9.40	4.71	0.05	9.08	5.14	0.09
	Berufsbezogene Motivation				0.02	0.11	0.87
		$R^2 = 0.39$			$\Delta R^2 = 0.00, p = 0.87$		
Erweiterungen mit Korrektur	Berufsbezogene Motivation	0.00	0.03	0.97			
		$R^2 = 0.00$					
Erweiterungen ohne Korrektur	Abschluss: Fachhochschule	9.71	3.67	0.01	8.71	4.02	0.04
	Position: stellv. Leitung	8.09	2.91	0.01	8.40	2.97	0.01
	Berufsbezogene Motivation				-0.04	0.07	0.53
		$R^2 = 0.25$			$\Delta R^2 = 0.01, p = 0.53$		
Anknüpfen	Feedbackgeber	-3.67	1.84	0.05	-3.95	1.96	0.05
	Position: stellv. Leitung	5.09	2.18	0.03	5.37	2.29	0.02
	Berufsbezogene Motivation				-0.02	0.05	0.66
		$R^2 = 0.19$			$\Delta R^2 = 0.01, p = 0.66$		

Tabelle 50: Kennwerte der hierarchischen Regressionen zur Vorhersage der sprachlichen Anregung durch das Reflexionsniveau

	Prädiktoren	Modell 1			Modell 2		
		B	SE B	p	B	SE B	p
Language Modeling	Language Modeling T1	-0.40	0.12	<0.01	-0.41	0.12	<0.01
	Abschluss: Kinderpflege	-0.72	0.32	0.03	-0.72	0.32	0.03
	Reflexionsniveau				-0.07	0.22	0.77
		$R^2 = 0.34$			$\Delta R^2 = 0.00, p = 0.78$		
Geschlossene Fragen mit AM	Geschlossene Fragen mit AM_T1	-0.44	0.19	0.03	-0.44	0.20	0.03
	Abschluss: Fachhochschule	11.89	5.61	0.04	11.92	5.69	0.04
	Reflexionsniveau				-0.35	3.56	0.92
		$R^2 = 0.22$			$\Delta R^2 = 0.00, p = 0.92$		
Ja/Nein-Fragen mit AM	Ja/Nein-Fragen mit AM_T1	-0.68	0.10	<0.01	-0.68	0.10	<0.01
	Position: FK ohne Gruppenleitung	9.25	4.67	0.06	8.68	4.68	0.07
	Reflexionsniveau				3.53	3.13	0.27
		$R^2 = 0.58$			$\Delta R^2 = 0.01, p = 0.28$		
Offene Fragen mit AM	Position: Leitung	21.69	6.63	<0.01	21.31	6.70	<0.01
	Reflexionsniveau				2.06	2.98	0.49
		$R^2 = 0.22$			$\Delta R^2 = 0.01, p = 0.49$		
Offene Fragen ohne AM	Position: Gruppenleitung	-2.95	1.44	0.05	-2.95	1.37	0.04
	Reflexionsniveau				2.36	1.06	0.03
		$R^2 = 0.1$			$\Delta R^2 = 0.11, p = 0.03$		
Erzählaufforderungen mit AM	Position: stellv. Leitung	15.86	5.06	<0.01	15.76	5.19	<0.01
	Reflexionsniveau				-0.45	3.66	0.90
		$R^2 = 0.21$			$\Delta R^2 = 0.00, p = 0.90$		
Wiederholung mit Korrektur	Wiederholung mit Korrektur_T1	-0.68	0.33	0.05	-0.68	0.32	0.04
	Reflexionsniveau				-0.53	0.86	0.54
		$R^2 = 0.09$			$\Delta R^2 = 0.01, p = 0.54$		
Wiederholung ohne Korrektur	Wiederh. ohne Korrektur_T1	-0.43	0.17	0.02	-0.45	0.18	0.02
	Feedbackgeber	7.70	3.91	0.06	7.28	4.07	0.08
	Abschluss: Kinderpflege	-3.13	5.13	0.55	-2.72	5.28	0.61
	Position: stellv. Leitung	9.40	4.71	0.05	9.78	4.85	0.05
	Reflexionsniveau				1.50	3.50	0.67
		$R^2 = 0.39$			$\Delta R^2 = 0.00, p = 0.67$		
Erweiterungen mit Korrektur	Reflexionsniveau	-0.23	0.81	0.78			
		$R^2 = 0.00$					
Erweiterungen ohne Korrektur	Abschluss: Fachhochschule	9.71	3.67	0.01	9.65	3.72	0.01
	Position: stellv. Leitung	8.09	2.91	0.01	8.28	2.97	0.01
	Reflexionsniveau				0.94	2.06	0.65
		$R^2 = 0.25$			$\Delta R^2 = 0.00, p = 0.65$		
Anknüpfen	Feedbackgeber	-3.67	1.84	0.05	-2.91	1.84	0.12
	Position: stellv. Leitung	5.09	2.18	0.03	4.41	2.15	0.05
	Reflexionsniveau				-2.76	1.55	0.08
		$R^2 = 0.19$			$\Delta R^2 = 0.07, p = 0.08$		

Tabelle 51: Kennwerte der hierarchischen Regressionen zur Vorhersage der sprachlichen Anregung durch die Skalen des "Questionnaire for Training Evaluation"

	Prädiktoren	Modell 1			Modell 2		
		B	SE B	p	B	SE B	p
Language Modeling	Language Modeling T1	-0.46	0.14	<0.01	-0.41	0.15	0.01
	Abschluss: Kinderpflege	-0.56	0.35	0.12	-0.88	0.39	0.03
	Q4TE_Satisfaction				-0.16	0.28	0.57
	Q4TE_Utility				0.32	0.17	0.08
	Q4TE_Knowledge				-0.11	0.08	0.21
	Q4TE_Application to practice				-0.03	0.15	0.87
	Q4TE_Individual organiz. results				-0.07	0.11	0.53
		$R^2 = 0.33$			$\Delta R^2 = 0.11, p = 0.34$		
Geschlossene Fragen mit AM	Geschlossene Fragen mit AM_T1	-0.43	0.21	0.05	-0.42	0.22	0.07
	Abschluss: Fachhochschule	12.28	5.93	0.05	13.01	6.28	0.05
	Q4TE_Satisfaction				-1.66	4.62	0.72
	Q4TE_Utility				-0.30	2.77	0.92
	Q4TE_Knowledge				0.03	1.36	0.98
	Q4TE_Application to practice				3.46	2.55	0.19
	Q4TE_Individual organiz. results				-1.67	1.78	0.36
		$R^2 = 0.23$			$\Delta R^2 = 0.07, p = 0.73$		
Ja/Nein-Fragen mit AM	Ja/Nein-Fragen mit AM_T1	-0.68	0.11	<0.01	-0.72	0.14	<0.01
	Position: FK ohne Gruppenleitung	8.80	4.76	0.07	9.34	5.03	0.07
	Q4TE_Satisfaction				5.66	4.58	0.23
	Q4TE_Utility				2.06	2.64	0.44
	Q4TE_Knowledge				0.38	1.18	0.75
	Q4TE_Application to practice				-2.63	2.48	0.30
	Q4TE_Individual organiz. results				0.45	1.50	0.77
		$R^2 = 0.60$			$\Delta R^2 = 0.05, p = 0.59$		
Offene Fragen mit AM	Position: Leitung	21.74	6.77	<0.01	22.75	7.05	<0.01
	Q4TE_Satisfaction				0.54	3.78	0.89
	Q4TE_Utility				2.60	2.49	0.31
	Q4TE_Knowledge				0.80	1.11	0.48
	Q4TE_Application to practice				0.51	2.17	0.81
	Q4TE_Individual organiz. results				-3.10	1.40	0.04
		$R^2 = 0.24$			$\Delta R^2 = 0.12, p = 0.42$		
Offene Fragen ohne AM	Position: Gruppenleitung	-2.89	1.66	0.09	-1.80	2.37	0.45
	Q4TE_Satisfaction				-0.60	1.42	0.68
	Q4TE_Utility				-0.68	1.09	0.54
	Q4TE_Knowledge				0.36	0.42	0.39
	Q4TE_Application to practice				0.05	0.83	0.95
	Q4TE_Individual organiz. results				0.72	0.53	0.19
		$R^2 = 0.08$			$\Delta R^2 = 0.15, p = 0.37$		
Erzählaufforderungen mit AM <sup>1</sup>	Position: stellv. Leitung	19.59	-0.12	0.07	16.72	-0.89	0.06
	Q4TE_Satisfaction				-0.73	-0.79	0.85
	Q4TE_Utility				-0.25	-0.51	0.93

		Modell 1			Modell 2		
	Prädiktoren	B	SE B	p	B	SE B	p
	Q4TE_Knowledge				1.42	-0.29	0.29
	Q4TE_Application to practice				4.21	-0.25	0.17
	Q4TE_Individual organiz. results				-4.68	0.92	0.09
		$R^2 = 0.29$			$\Delta R^2 = 0.22, p = 0.05$		
Wiederholung mit Korrektur	Wiederholung mit Korrektur_T1	-0.82	0.23	<0.01	-0.85	0.23	<0.01
	Q4TE_Application to practice				0.40	0.32	0.22
	Q4TE_Individual organiz. results				-0.22	0.24	0.37
		$R^2 = 0.26$			$\Delta R^2 = 0.03, p = 0.46$		
Wiederholung ohne Korrektur	Wiederh. ohne Korrektur_T1	-0.44	0.17	0.02	-0.35	0.21	0.10
	Feedbackgeber	6.63	3.91	0.10	6.06	4.21	0.16
	Abschluss: Kinderpflege	-4.26	5.23	0.42	-9.35	6.61	0.17
	Position: stellv. Leitung	6.76	4.79	0.17	4.88	5.45	0.38
	Q4TE_Satisfaction				-2.35	4.27	0.59
	Q4TE_Utility				3.90	2.90	0.19
	Q4TE_Knowledge				-0.24	1.23	0.85
	Q4TE_Application to practice				-0.73	2.37	0.76
	Q4TE_Individual organiz. results				-0.87	1.56	0.58
		$R^2 = 0.39$			$\Delta R^2 = 0.05, p = 0.81$		
Erweiterungen mit Korrektur	Q4TE_Satisfaction	0.83	1.15	0.47			
	Q4TE_Utility	-0.34	0.70	0.63			
	Q4TE_Knowledge	0.27	0.32	0.40			
	Q4TE_Application to practice	0.40	0.64	0.54			
	Q4TE_Individual organiz. results	-0.20	0.44	0.65			
		$R^2 = 0.07$					
Erweiterungen ohne Korrektur	Abschluss: Fachhochschule	9.67	3.84	0.02	10.51	3.63	0.01
	Position: stellv. Leitung	6.65	3.20	0.05	6.22	3.12	0.06
	Q4TE_Satisfaction				-6.33	2.37	0.01
	Q4TE_Utility				-0.97	1.60	0.55
	Q4TE_Knowledge				-0.61	0.70	0.40
	Q4TE_Application to practice				2.85	1.39	0.05
	Q4TE_Individual organiz. results				0.65	0.91	0.49
		$R^2 = 0.22$			$\Delta R^2 = 0.23, p = 0.08$		
Anknüpfen <sup>1</sup>	Feedbackgeber	-3.54	-0.01	0.12	-4.09	-0.33	0.09
	Position: stellv. Leitung	4.34	0.05	0.08	4.63	0.40	0.12
	Q4TE_Satisfaction				-3.50	0.09	0.07
	Q4TE_Utility				-0.71	-0.37	0.62
	Q4TE_Knowledge				-1.33	-0.20	0.06
	Q4TE_Application to practice				1.64	-0.26	0.18
	Q4TE_Individual organiz. results				1.09	0.36	0.15
		$R^2 = 0.16$			$\Delta R^2 = 0.24, p = 0.10$		

Anmerkung: <sup>1</sup> Aufgrund von Heteroskedastizität erfolgte eine korrigierte Signifikanzprüfung mithilfe des Bootstrap-Verfahrens.

Tabelle 52: Kennwerte der hierarchischen Regressionen zur Vorhersage der sprachlichen Anregung durch technische Probleme während des Feedbacks

	Prädiktoren	Modell 1			Modell 2		
		B	SE B	p	B	SE B	p
Language Modeling	Language Modeling T1	-0.40	0.12	<0.01	-0.39	0.12	<0.01
	Abschluss: Kinderpflege	-0.72	0.32	0.03	-0.70	0.32	0.04
	Technische Probleme				-0.24	0.28	0.40
		$R^2 = 0.34$			$\Delta R^2 = 0.01, p = 0.40$		
Geschlossene Fragen mit AM	Geschlossene Fragen mit AM_T1	-0.44	0.19	0.03	-0.46	0.20	0.03
	Abschluss: Fachhochschule	11.89	5.61	0.04	12.11	5.67	0.04
	Technische Probleme				2.53	4.54	0.58
		$R^2 = 0.22$			$\Delta R^2 = 0.01, p = 0.58$		
Ja/Nein-Fragen mit AM	Ja/Nein-Fragen mit AM_T1	-0.68	0.10	<0.01	-0.68	0.11	<0.01
	Position: FK ohne Gruppenleitung	9.25	4.67	0.06	9.24	4.78	0.06
	Technische Probleme				0.12	4.03	0.98
		$R^2 = 0.58$			$\Delta R^2 = 0.00, p = 0.98$		
Offene Fragen mit AM	Position: Leitung	21.69	6.63	<0.01	20.10	6.70	0.01
	Technische Probleme				-4.62	3.70	0.22
		$R^2 = 0.22$			$\Delta R^2 = 0.03, p = 0.22$		
Offene Fragen ohne AM	Position: Gruppenleitung	-2.95	1.44	0.05	-2.96	1.46	0.05
	Technische Probleme				0.24	1.40	0.87
		$R^2 = 0.1$			$\Delta R^2 = 0.00, p = 0.87$		
Erzählaufforderungen mit AM	Position: stellv. Leitung	15.86	5.06	<0.01	15.79	5.09	<0.01
	Technische Probleme				3.19	4.46	0.48
		$R^2 = 0.21$			$\Delta R^2 = 0.01, p = 0.48$		
Wiederholung mit Korrektur	Wiederholung mit Korrektur_T1	-0.64	0.32	0.05	-0.51	0.33	0.13
	Technische Probleme				-1.47	1.11	0.19
		$R^2 = 0.08$			$\Delta R^2 = 0.04, p = 0.19$		
Wiederholung ohne Korrektur	Wiederh. ohne Korrektur_T1	-0.43	0.17	0.02	-0.45	0.17	0.02
	Feedbackgeber	7.70	3.91	0.06	8.85	4.31	0.05
	Abschluss: Kinderpflege	-3.13	5.13	0.55	-3.17	5.18	0.54
	Position: stellv. Leitung	9.40	4.71	0.05	9.12	4.77	0.06
	Technische Probleme				2.93	4.43	0.51
		$R^2 = 0.39$			$\Delta R^2 = 0.01, p = 0.51$		
Erweiterungen mit Korrektur	Technische Probleme	-0.12	1.03	0.91			
		$R^2 = 0.00$					
Erweiterungen ohne Korrektur	Abschluss: Fachhochschule	9.71	3.67	0.01	11.05	3.63	<0.01
	Position: stellv. Leitung	8.09	2.91	0.01	8.20	2.82	0.01
	Technische Probleme				4.59	2.46	0.07
		$R^2 = 0.25$			$\Delta R^2 = 0.07, p = 0.07$		
Anknüpfen	Feedbackgeber	-3.67	1.84	0.05	-2.91	2.04	0.16
	Position: stellv. Leitung	5.09	2.18	0.03	4.96	2.19	0.03
	Technische Probleme				1.86	2.10	0.38
		$R^2 = 0.19$			$\Delta R^2 = 0.02, p = 0.38$		